THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/1108

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

Self-Assessment in Religious Education

by

Nigel Peter Michell Fancourt

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Educational Studies

University of Warwick, Institute of Education

April 2008

Contents

Tables and figures ix

Acknowledgements and declaration x

Abstract xi

Glossary xii

1. A professional dilemma: defining the problem 1

2. Self-assessment as assessment for learning 7

    2.1. Self-assessment in context 8

    2.2. Challenging the ascendancy of summative assessment 10

    2.3. The benefits of formative assessment 15

        2.3.1. Questioning 17

        2.3.2. Feedback 17

        2.3.3. Sharing criteria 19

        2.3.4. Self-assessment 20

    2.4. Criticisms of Black and Wiliam 24

        2.4.1. Genuine success? 24

        2.4.2. Pupil identity 25

        2.4.3. Other relevant research 27

        2.4.4. A place for summative assessment 27

        2.4.5. A narrow view of feedback 28

    2.5. Formative assessment in the classroom 30

    2.6. Contrasting theories of assessment and learning 32

    2.7. Framing a general question on self-assessment as assessment for learning 33

2.8. Conclusion 34

3. Literature review: pedagogy and policy in religious education 36

3.1. 'Learning about religion' 40

3.1.1. Learning about religion - the Model Syllabuses 42

3.1.2. Learning about religion - Non-statutory guidance 48

3.1.3. Learning about religion - the Framework 51

3.2. 'Learning from religion' 54

3.2.1. Learning from religion - the Model Syllabuses 57

3.2.2. Learning from religion - Non-statutory guidance 59

3.2.3. Learning from religion - the Framework 65

3.3. Attitudes and values in religious education 67

3.4. Self-assessment in religious education 71

3.5. Policy and pedagogy in religious education: a general question 74

3.6. Conclusion 75

4. The virtues of practitioner research: methodological issues 77

4.1. Virtue theory as a research paradigm 78

4.1.1. What are research paradigms? 79

4.1.2. The strengths of virtue theory as a research paradigm 82

4.1.3. Balancing two practices 85

4.2. Practitioner research 87

4.2.1. Emancipation or professional development? 88

4.2.2. Questions of quality 93

4.3. Conclusion 98

5. Structuring the research: research questions and design 100

5.1. Qualitative or quantitative? 101

5.2. From general questions to research questions 103

5.3. Research design for practitioner research 108

5.3.1. The action steps 108

5.3.2. Data collection 112

5.3.3. Data analysis 116

5.3.4. The lack of action cycles 118

5.4. Ethical justification 120

5.5. Conclusion 121

6. Teaching *and* researching: a review of the research process 123

6.1. Lord Williams's School 123

6.2. The pilot study 124

6.2.1. The pilot study : the class 125

6.2.2. The pilot study : implementing the action steps 126

6.2.3. The pilot study: data collection and analysis 127

6.2.4. The pilot study : design modification 128

6.3. The main research 129

6.3.1. The main research : the class 129

6.3.2. The main research: the action steps 130

6.3.3. The main research: data collection 131

6.4. Teacher researcher 133

6.4.1. Influencing pupils 133

6.4.2. Reactivity 135

6.4.3. 'Teacherly' reactions 135

6.4.4. Exasperation 136

6.4.5. Teacher researcher: conclusion 137

6.5. Pupil relationships 137

6.5.1. Social activity 138

6.5.2. Games 138

6.5.3. Evolving collective thought 139

6.5.4. Pupils as co-interviewers 140

6.5.5. Pupil relationships: conclusion 142

6.6. Objectivity in data analysis and theory generation 143

6.7. Conclusion 144

7. From techniques to themes: analyzing the pilot study 146

7.1. The pilot research questions 147

7.2. The pilot questionnaires and the interview coding 147

7.3. Pupils' sense of progress 148

7.4. Pupils' understanding of self-assessment 149

7.5. Answering the research questions 150

7.5.1. Traffic-lighting 150

7.5.2. Formative use of summative assessment 153

7.5.3. Class assessment criteria 154

7.5.4. Discussing tolerance 156

7.6. Theoretical implications 159

7.6.1. Differentiation and variety 159

7.6.2. The use of levels 161

7.6.3. Assessment as a social activity 164

7.6.4. Sharing criteria 166

7.7. Discussion of themes in the pilot study 168

7.8. Conclusion 169

8. Self-assessment and pupil identity: analyzing the main research 171

8.1. Pupils' understanding of self-assessment 172

8.2. Pupils' views of summative processes 173

8.2.1. Summative processes: ego-focused responses 175

8.2.2. Summative processes: task-focused responses 179

8.2.3. Summative processes: conclusion 180

8.3. Pupils' views of self-assessment 180

8.3.1 Task or ego-focus in self-assessment 182

8.3.2 The teacher and the pupils 186

8.3.3 Understanding the criteria 189

8.3.4 Socio-cultural dynamics 192

8.3.5 Pupils' views of self-assessment: conclusion 197

8.4. Pupils' views of self-assessment using levels 198

8.4.1. Self-assessment using levels: ego-focus 199

8.4.2. Self-assessment using levels: task-focus 205

8.5. Answering the research question on assessment processes 206

8.6. Conclusion 209

9. Reflection in religious education: analysing the main research 211

9.1. No value in religious education 212

9.2. Pupils' views of 'learning about religion' 214

9.3. Pupils' views of 'learning from religion' 217

9.3.1. Constructive criticism 218

9.3.2. Edification 220

9.3.3. Expression 222

9.3.4. Implications for the current attainment targets 224

9.4. Self-assessment of the development of values 225

9.4.1. Tolerance 226

9.4.2. Respect 230

9.5. Progress in religious education 234

9.6. Conclusion 237

10. Self-assessment in religious education: three pupils' voices 239

10.1. Pupil's experiences of assessment in religious education 240

10.2. Aaron 241

10.3. Jenni 249

10.4. Kevin 256

10.5. Assessment careers: discussion 267

10.6. Conclusion 271

11. Reflexive self-assessment: assessment and pedagogy in religious education 273

11.1. From the research questions to the general aim 273

11.2. Reflexive self-assessment 275

11.3. Comparing models of self-assessment 280

11.3.1. Understanding the goals 281

11.3.2. Sharing the criteria 284

11.3.3. Current 'attainment' 286

11.3.4. Closing the gap 287

11.4. Reflexive self-assessment and the use of levels 289

11.5. Learning autonomy and its limits 291

11.6. The wider context of reflexive self-assessment 296

11.7. Conclusion 298

12. Reflexive self-assessment in religious education 300

12.1. Reflexive self-assessment and the pedagogy of religious education 300

12.1.1. Intellect and values 301

12.1.2. Interpretive approaches and action research 303

12.2. The attainment targets revisited 307

12.2.1. The context of text production 308

12.2.2. Problems with the Framework 310

12.2.3. Redrafting the Framework 313

12.3. Conclusion 314

13. Conclusion: reflections on the research 316

13.1. Reflecting on the research process 316

13.2. Reflecting on the practices of researching and teaching 317

13.3. Reflecting on teaching while researching 320

13.4. Future developments 323

13.5. Contribution to educational research 326

14. Appendices 330

14.1. Module traffic-light sheet 330

14.2. Example of written self-assessment comments 331

14.3.	Level-path sheet	332
14.4.	Timetable for teaching and data collection in the main research	333
14.5.	Draft pilot questionnaire	335
14.6.	Transcription conventions	336
14.7.	Pilot study codes	337
14.8.	Main research codes	338
14.9.	Main research interview timetable	342
14.10.	First pilot questionnaire	343
14.11.	Second pilot questionnaire	344
14.12.	Ten principles for assessment in religious education	345

15. References 348

Tables and Figures

Figure 1: Levels of reflection in religious education 73

Table 1: Question structure in the research 107

Figure 2: The components of reflexive self-assessment 278

Acknowledgements

This research would not have been possible without the help and support of many people and organizations. Firstly, I am very grateful for funding from, in particular, Culham Educational Foundation, and from the Institute of Education at Warwick; I would also like to thank the Farmington Trust for funding the sabbatical term which led to this research. Secondly, my supervisors, Bob Jackson and Val Brooks, have been more than generous in their discerning, critical support throughout the entire project, encouraging and developing my research. Thirdly, there are many people who have offered comments and suggestions along the way, notably in the WRERU doctoral research group and the REDCo 'community of practice'. In particular I would like to thank Lat Blaylock, Julia Ipgrave, Rachael Jackson, Mary Hayward, Eleanor Nesbitt, Kevin O'Grady, and Barbara Wintersgill. Fourthly, this would not be possible without the co-operation and support of colleagues and pupils at Lord Williams's School.

Finally, however, I am forever indebted to Amanda, for continuous, reliable, meticulous support throughout the five years, and to Eleanor for stopping me taking it too seriously.

Declaration

This thesis is entirely my work. None of it has appeared in any form before, except where clearly referenced extracts from published articles are quoted. None of it has been submitted for a degree at another university.

Abstract

This research investigates the nature of pupil self-assessment in religious education. It considers the implications of theories of self-assessment as assessment for learning for self-reflection in pedagogies of pluralistic religious education, and *vice versa*.

Assessment for learning: Research on assessment has claimed that self-assessment is essential in formative assessment, to combat the negative effects of summative assessment. Other recent research has considered the situated nature of classroom practice. How would these classroom factors affect self-assessment in RE?

Policy and pedagogy in religious education: The history of the current policy documents is analysed using policy scholarship, and the tension is revealed between measurable intellectual skills and a wider understanding of the place of religious education in developing tolerance and respect, both in the England and Wales, and internationally. Are policy and assessment properly aligned?

Practitioner research: Virtue theory is developed as a research paradigm for practitioner research for professional development. Rigour is established through a reflexive use of qualitative, largely ethnographic methods, especially group interviews. Analysis includes consideration of pupils' assessment careers.

***Reflexive* self-assessment:** As a result of analyzing the data on assessment and religious education an original form of self-assessment is proposed. *Reflexive* self-assessment is a subject-specific model of self-assessment, linked to interpretive approaches. This harmonizes classroom self-assessment of both intellectual skills and intercultural values. The classroom conditions necessary to allow it to develop are examined. The implications of this for theories of self-assessment, learning autonomy and current policies of religious education are considered.

Finally, the research is reviewed, notably the implications for researching and teaching, and future developments. The quality of the research is defended, in terms of significance, originality and rigour.

Glossary

AAIA: Association for Achievement and Improvement through Assessment

AREIAC: Association of Religious Education Inspectors, Advisors and Consultants

Assessment for learning: Term currently used in national documents to describe formative assessment

Assessment of learning: Term currently used in national documents to describe summative assessment

BERA: British Educational Research Council

DfES: Department for Education and Skills

ENRECA: European Network for Religious Education through Contextual Approaches

ESRC: Economic and Social Research Council

FARE Project: 'Forms of Assessment in Religious Education' project; a model of assessment in religious education, produced by researchers at Exeter University (see Copley *et al.* 1991), working with eight local authorities

The Framework: exemplar framework for local syllabuses, including assessment criteria (set out in QCA 2005)

The Guidelines: assessment guidelines, mirroring assessment criteria in the National Curriculum, based on the Model Syllabuses (set out in QCA 2000)

INSET: In service training; the term for teachers' professional development events, both in school and out of school

KMOFAP: King's Medway Oxfordshire Formative Assessment Project; a research project that investigated the implementation of assessment for learning in schools (see Black *et al.* 2003)

Learning about religion: The first attainment target in the Model Syllabuses, the Guidance and the Framework, covering pupils' knowledge and understanding of religious beliefs, practices and values

Learning from religion: The second attainment target in the Model Syllabuses, the Guidance and the Framework, covering pupils' evaluation of and reflection on religious beliefs, practices and values

Learning how to Learn Project: a research project investigating the wider conditions for assessment for learning (see James *et al.* 2007)

Model Syllabuses: Exemplar syllabuses for religious education, for local authorities and teacher training (set out in SCAA 1994)

OECD: Organization for Economic Co-operation and Development

OFSTED: Office for Standards in Education, the English schools' inspection agency

OSCE: Organization for Security and Co-operation in Europe

QCA: Qualifications and Curriculum Authority, the current government agency responsible for qualifications and curriculum matters

REDCo Project: 'Religion in education; a contribution to Dialogue or a factor of Conflict in transforming societies of European countries' project; an EU-funded European comparative research programme into the place of religion in education (see Jackson *et al.* 2007)

SACRE: Standing Advisory Committee on Religious Education; the local authority committee of representatives from different religious traditions that sets the local agreed syllabus, under the Education Reform Act 1988

SCAA: School Curriculum and Assessment Authority, the government agency responsible for school education in the early 1990s (replaced by QCA)

Westhill Project: a model of assessment in religious education, produced by researchers at Westhill College, Birmingham (see Westhill College 1989, Westhill College 1991), working with twenty-nine local authorities

Chapter One

A professional dilemma: defining the problem

Like many teachers, my experience of whole school INSET sessions is varied. On one hand, they can be frustrating, being given obvious or irrelevant advice. These are very frustrating when one has lessons to plan, books to mark, parents to contact, trips to organise and reports to write. On the other hand, they can sometimes provide genuinely useful insights, which affect one's practice thereafter.

A few years ago my school organised an INSET day; it was a mixture of both the obvious and the inspiring. There was a range of presenters; some were from the school - maths, English and science teachers - and one was an academic with whom they were working on a research project – Dylan Wiliam. The research was on formative assessment, and the teachers had been trialling different techniques in their lessons to see what difference these made to their teaching; they were quite excited about it. However, some of the advice was, to me, obvious – such as not to give marks or grades on work but comments instead. Other aspects of it, especially the summary of research on the effects on pupils of grades and marks was insightful.

This was part of the King's Medway Oxfordshire Formative Assessment Project ("KMOFAP"). Researchers from King's College in London were investigating how formative assessment, or assessment for learning, could be embedded in schools. Two of them, Paul Black and Dylan Wiliam, had earlier reviewed the voluminous research evidence on its effectiveness (Black and Wiliam 1998a); they now wanted to consider how research could be transformed into practice (Black and William 1998b). They had set up links with schools on the Medway in Kent and in

Oxfordshire to investigate this more fully (described in Black *et al.* 2002, Black *et al.* 2003). This research was not simply about the effects on pupil attainment, but rather on what had to happen in schools to bring about these effects, in terms of teachers' practices, school management and culture.

I pondered some of these ideas in the light of the challenge of being a teacher of religious education. I wanted to present the pupils with a vibrant subject, so that they would value their lessons. I tried constantly to teach dynamic high-impact lessons that involved pupils through role-plays, artwork and discussions and therefore, I hoped, motivate them. I believed that if they were busy and enjoying themselves then they would make progress. But often this did not seem to be the case; they seemed to enjoy the 'fun' activities, but their work seemed to lack effort or depth. Thus, I also felt that I had to keep up with the marking of their work, in order to give them the feedback that they needed to improve. In my own education, I was very grateful to teachers who had given me constructive feedback, and considered this an essential part of professionalism.

What should I tell my pupils to do in the feedback? In terms of basic knowledge and understanding, this was fairly straightforward. It was also possible to encourage them to express their ideas more clearly and fully; but what advice should I give about more complicated processes, such as reflection and empathy? If pupils could not empathise being someone else then how could I describe this to them?

Moreover, even when I gave feedback, pupils did not seem to understand what the purpose of the subject was, or why they were doing it. They seemed to enjoy the lessons, and understood the feedback, but still maintained that they did like religious education, or that the lessons were not really religious education.

Obviously, any approach that would encourage their involvement and motivation, reduce the marking load, *and* give them a sense of educational development would be worthwhile.

The first formal stage of reflecting on my practice led me to a review of the place of empathy in current pedagogy (Fancourt 2003a). This was through a Farmington fellowship, which allowed me to take a sabbatical term to research the topic. This was a theoretical review, though it had a practical goal. Empathy is often presented as an important quality that pupils can develop through religious education, but it is hard to define, and also the definitions differ: is it emotional or is it cognitive? Is it the same as sympathy? Is it the same as tolerance? How would a teacher know if a pupil had developed it, and what feedback would one give? This sabbatical term allowed me to review some of the literature on empathy in religious education as well as the literature and research on assessment for learning. Aspects of this were published in a teachers' journal (Fancourt 2003b).

Towards the end of the sabbatical, I began to focus on a related problem. Even if the feedback was effective, even if I could give the pupils good advice, the time spent marking and writing the comments was time that could otherwise be spent planning those exciting lessons. It was impossible to do both. I could give better advice, but would the lessons now be dull and boring for the pupils? What was needed was a more radical solution. I needed a way for the pupils to grasp what their learning was about. If they understood what they were supposed to be achieving, then they could take more control of their progress. This is self-assessment, as described by the KMOFAP research.

However, this is tantalisingly similar to ideas about how pupils should reflect personally on their learning in religious education, and not simply gain knowledge

of different religions. National guidance on religious education had highlighted the need for pupils to 'learn from' religion and not just 'learn about' it (SCAA 1994). Furthermore, the locally agreed syllabus for Oxfordshire at Key Stage 3 (11 to 14 years old) emphasised the need for pupils to 'develop their critical appraisal of some of the aspects of religion' and to 'reflect and build on their own personal experience' (Oxfordshire County Council 1992, 46).

I therefore increasingly focused on the connection between these two processes. Was self-assessment the same as or different from reflection in religious education? Did religious education automatically include the key elements of formative assessment – was this another example of being told the banal? Or did this new approach to assessment have something more focused to offer religious education? Indeed, does religious education have a contribution to make to generic theories of self assessment? I discussed this inter-relationship in an earlier article (Fancourt 2005a).

The heart of the matter, however, is one general issue:

- What is the nature of self-assessment in religious education?

My research sets out to investigate this issue, and is organised into four main sections.

Firstly, the relevant literature is reviewed. There is however almost no research on the inter-relationship between ideas of formative assessment and the pedagogy of religious education. The next two chapters therefore review relevant literature and research separately: chapter two reviews work on assessment for learning, especially self-assessment; the third chapter reviews writing on religious

education. These chapters each suggest a *general* question.

In the next three chapters, methodological issues are discussed. Chapter four tackles the ethical and epistemological underpinnings of practitioner research methodology, arguing for the comprehensive use of virtue theory. Chapter five then deals with the detail of method and research design, including the pilot study, in which the *general* questions that emerged from the literature review are narrowed down to *research* questions, which the research can seek to answer. The answers to the research questions will then inform the general questions. The sixth chapter reviews the methodological issues that arose in the conduct of the research, especially the trustworthiness of the interviews in the context of practitioner research.

The third section contains the analysis of the data. Chapter seven explores the substantive issues that emerged from the pilot study. Chapter eight considers generic assessment issues from the main research, and the next chapter reviews the implications for pedagogy and policy in religious education, particularly the tensions between knowledge and reflection, and attitudes and values. The tenth chapter looks at how these two sets of issues are entangled in the experiences of individual pupils. These chapters relate the data to the research questions.

The last section contains two chapters, which consider the implications of the research for the *general* questions; one develops the implications of the analysis, and argues for a specific understanding of self-assessment in religious education as *reflexive* self-assessment, which incorporates the development of both intellect and values. The next, and penultimate, chapter explores some of the wider implications of this theory in the development of pedagogy and policy in religious education.

Each chapter concludes with a review of its contribution to the research, for instance in terms of its significance academically or practically, its originality of approach, or its contribution to issues of methodological rigour. In the last chapter, the research is reflected upon and then evaluated, linking the various elements together.

For the avoidance of doubt, it is important to stress that this research has no connection with any of Black and Wiliam's research, or KMOFAP, or other related research with which my school has been involved. Nevertheless, their research provided the starting point for mine, so it is only logical to move on to a review of their ideas.

Chapter Two

Self-assessment as assessment for learning

In order to make sense of the nature of self-assessment in religious education, the relevant research literature on self-assessment in religious education should be explored, but the problem is the general lack of it. The only *direct* work is my own review in a paper entitled 'Challenges for self-assessment in religious education' (Fancourt 2005a). Two major research-based practical guides to assessment in religious education both accorded an important role to self-assessment: the Westhill Project, from Westhill College in Birmingham (Westhill College 1989, 1991); the 'Forms of Assessment in Religious Education' Project ('FARE'), from the University of Exeter, which was the result of collaborative action research (Copley *et al.* 1991, 17). Both are considered in chapter three. It also features in other practical guidance (e.g. OFSTED 2003, Draycott 2006).

To understand the issue, the wider field of research literature, pedagogy, policy documents and professional guides needs to be reviewed. Firstly, in this chapter, the generic literature and research on the context of self-assessment in wider debates about formative and summative assessment is considered. Secondly, in the following chapter, the background literature on pedagogy and policy in religious education is analysed, especially current attainment targets, and then the inter-relationship between these two areas of research is explored. For each of these two areas, a general question is identified, which will inform the research.

2.1 Self-assessment in context

The first area for review is that of generic approaches to self-assessment. The focus is on this as a part of classroom assessment, i.e. what teachers and pupils do in the course of their day-to-day practice. However, the moment that one talks about assessment, a whole set of general educational issues come into play, such as marking, grades, tests and examinations. Assessment is a controversial issue; it 'is not a simple or innocent term' (Black *et al.* 2003, 1); it casts a 'long, dark shadow' (Bowe *et al.* 1992, 101). Indeed, the advent of a test-dominated approach to education in the 1990s led Grimmitt to argue that 'RE had fallen victim to a technicist and standards-related political ideology of education' (Grimmitt 2000b, 7); Copley claims that 'danger is not too strong a word for it' (Copley 2005, 16, see also Torrance 1997). The stakes are high because what one commentator sees as being a positive development another will see as positively harmful.

What is assessment? Assessment can be defined as 'a procedure for eliciting evidence that can assist in educational decision making' (Wiliam 1994, 5), and it can be considered in a number of ways, including type, purpose and identity of the assessor. The most important distinction however is between two assessment paradigms; a paradigm is 'a set of interrelated concepts which provide a framework within which we see and understand a particular problem or activity' (Gipps 1994, 1, see also Aikenhead 1997). The first is a summative assessment paradigm, also called 'assessment of learning'; this is a concern with the *measurement* of attainment, notably in public examinations, e.g. GCSEs

and A levels in England and Wales. Further, it can apply to the use of tests in schools, and the giving of grades, marks or levels on a piece of work.

By contrast, formative assessment or 'assessment for learning' is aimed at helping the pupil's education progress. It is often associated with the idea of feedback, and particularly feed-forward. Assessment for learning can be seen as focused on *learning*, rather than on measuring. It draws attention to the root of the word 'assessment' in the Latin word '*assidere*', which can mean 'to sit by', 'to attend', or 'to assist' (Brooks 2002, 63), though it can also mean 'to besiege', according to Collin's Latin Dictionary. As Buckle observes, 'the term "formative" draws attention to the fact that the student's knowledge and understanding are still being developed' (Buckle 1990, 61).

These are other purposes of assessment. It can also be approached diagnostically, as 'a judgement which precisely indicates what aspects of learning a pupil has mastered, and what aspects need attention' (Copley *et al.* 1991, 83), often by trying to identify the underlying causes of the state of a pupil's attainment. This approach has dominated thinking in religious education; the primary description of assessment by John Keast, for many years the subject advisor at the Qualifications and Curriculum Agency, was as 'an integral part of the cycle of *planning* effective teaching and learning in any subject' (Keast 2003a, 56) (emphasis added). It can also be *evaluative*, as 'a judgement about the extent to which the learning opportunities offered are balanced and effective in achieving the objectives' (Copley *et al.* 1991, 83), and thus used by the teacher to judge the effectiveness of their teaching (Brooks 2002, 37-38).

Some guides to assessment from the later 1980s and early 1990s discussed a range of other purposes, such as motivation, selection and certification (e.g. Riding and Butterfield 1990). Assessment can therefore be seen as serving a number of purposes, but the central tension has been seen as being between the two paradigms: assessment of learning and assessment for learning. Why do some educationalists and policy-makers want to introduce more summative assessment throughout schools, and why do others bitterly oppose them?

2.2 Challenging the ascendancy of summative assessment

As far as summative assessment is concerned, policymakers have often argued, particularly since the early 1990s, that the introduction of formal assessment in schools, especially examinations, has raised educational standards. Testing is said to provide clear evidence of the effectiveness of schools; for instance, Tony Blair argued that exam results were 'real and hard evidence' of improvements in school standards (BBC 2004). Other research supports this; for instance, Brooks argues that the widespread introduction of the General Certificate of Education for secondary modern schools in the 1960s made a considerable difference to the educational achievement of many pupils (Brooks 2006). In religious education, Keast argues that the introduction of the GCSE in 1988 'raised the awareness of assessment among teachers and sharpened considerably their thinking about the cycle of planning teaching and learning' (Keast 2003a, 57). Furthermore, OFSTED has claimed that standards of educational attainment in the subject have improved as a result of pupils being entered for the short course GCSE in religious education. It argues that pupils

were motivated by the fact that their studies were now accredited (OFSTED 2005).

However, there have been a number of criticisms of an approach to schooling and education that gave such a central place to summative assessment. Andrew Davis, for instance, outlines a number of complaints, including the injustice of holding schools to account on the basis of test results, assuming that improved test results equate with raised standards, and that teaching to the test distorts the curriculum (Davis 1999). This echoes Grimmitt's ire at 'technicist culture', quoted above (Grimmitt 2000b). Similarly, the Swiss educationalist Perrenoud suggests that schools have become like restaurants that continually report on how the meal is being prepared, rather than actually getting on with preparing it (Perrenoud 2005).

The concern here is not with issues of accountability or standards, because this research is neither focused on school effectiveness nor on examinations (see e.g. Torrance 1997). However, it is important to grasp the contested nature of assessment. Further, it is vital to see how this approach can become an unquestioned part of ordinary classroom practice; for many people, 'assessment' essentially means tests and exams.

This approach has also been the subject of much criticism by those who argue for a wholesale development of assessment for learning in classrooms. In the 1980s, the Task Group on Assessment and Testing, chaired by Paul Black, advocated the primacy of formative assessment, although this was rejected by

Margaret Thatcher as being 'teacher dominated and uncosted' (Thatcher 1993, 594, in Daugherty and Ecclestone 2006, 154). In the 1990s, various academics maintained the argument for formative assessment, e.g. Caroline Gipps' *Beyond Testing: Towards a Theory of Educational Assessment* (1994).

The most significant work in English was Black and Wiliam's review article 'Assessment and Classroom Learning' (Black and Wiliam 1998a), which provided an extensive review of existing research and synthesised theory with practice; it was described as 'influential' (OECD 2005, 22), 'seminal' (Tierney 2006, 239) and 'very comprehensive and carefully researched' (Biggs 1998, 103). The main body of the article was a review of two hundred and fifty pieces of largely quasi-experimental research (from an initial list of nearly seven hundred) mostly published after 1988. Their article did not refer to earlier research than this because it was reviewed in two earlier articles (Natriello 1987, Crooks 1988). From their review, they argued that:

- Classroom evaluation practices generally encourage superficial and rote learning, concentrating on recall of isolated details, usually items of knowledge which pupils soon forget.

- Teachers do not generally review the assessment questions that they use and do not discuss them critically with peers, so there is little reflection on what is being assessed.

- The grading function is over-emphasised and the learning function

under-emphasised.

- There is a tendency to use a norm-referenced rather than a criterion approach, which emphasises competition between pupils rather than personal improvement of each. The evidence is that with such practices the effect of feedback is to teach the weaker pupils that they lack ability, so that they are de-motivated and lose confidence in their own capacity to learn (Black and Wiliam 1998a, 17-18).

The last point highlights one thread that runs through much of this research: the effect on pupil self-esteem, self-efficacy and motivation of summative assessment in the classroom. Research has identified two broad patterns of self-belief among students (e.g., Gipps 1994, Torrance and Pryor 1998, Dweck 1999). On the one hand, there are task-focused pupils: those with high self-esteem as learners – what Dweck calls the 'mastery' pattern (Black and Wiliam 1998a, 49; Dweck 1999, 7). These pupils are prepared to tackle challenges in learning, are not afraid to make mistakes and take intellectual risks.

On the other hand, there are self-focused pupils, or what Dweck calls the 'helpless' pattern (Black and Wiliam 1998a, 49; Dweck 1999, 7). Pupils in this category are afraid of failure, dislike challenges and want to give 'safe' answers. Moreover, the distinction did not necessarily coincide with exam success: some students who were successful in exams were self-focused, attributing success to their intelligence. But they were then at a loss when they confronted failure, attributing it to a lack of some quality – memory or intelligence. On the other

hand some comparatively less successful task-focused students could be positive about their education, because they focused on their own improvement.

It is therefore argued that summative assessment processes tend to foster self-focused students. This is because the use of grades, marks or levels tends to encourage competition, or what is termed norm-referencing:

> Competition is central to norm referencing, which is a form of social comparison. This can lead to severe discouragement for students who have few academic successes in competition with their peers. It also discourages students from helping each other with their academic work and encourages the covering up of misunderstandings... It also tends to encourage students to attribute success and failure to ability, rather than to effort, which is especially harmful for low-achieving pupils (Gipps 1994, 41).

In summary, a summative approach, implemented at the level of a national policy for school improvement, can have an invidious effect on a pupil's education. This is not because some pupils do not succeed, but because the approach creates a climate in which other pupils become doomed to fail. This is also confirmed in Harlen and Deakin-Crick's research (2002), which showed that 'after the introduction of the National Curriculum tests in England, low-achieving pupils had lower self-esteem than higher-achieving students' and the gap between the low-achieving and the high-achieving pupils widened. Furthermore, the 'successful' succeed for the 'wrong' reasons in that they are not task-orientated, but simply adept at passing exams, and unable to deal with failure

when it arises. As they point out, 'a strong emphasis on testing produces students with a strong extrinsic orientation towards grades and social status' (Harlen and Deakin Crick 2002). Rather than teaching to the test, it is learning for the test; the problem is not new (Bicknell 1919, viii), but the extent of it is. These pupils are not independent, reflective, 'deep' learners, able to rise to the challenge of failure. Filer and Pollard argue that 'if the state schooling is to produce such learners then a rethink of the use of assessment is essential' (Filer and Pollard 2000, 158). Is there an alternative?

2.3 The benefits of formative assessment

Many passionately argue that formative assessment is the required 'rethink'. As a process that focuses on learning, assessment should be embedded in the needs of pupils rather than the needs of the auditors of school standards. Further, it should be able to tackle the issues of extrinsic motivation, surface learning and 'helplessness'. Pupils would focus on the task, not on an ineffable quality such as 'intelligence'. As I have earlier remarked 'it is about the use of assessment to enable pupils to know more, understand better, and respond more deeply' (Fancourt 2005a, 116).

Black and Wiliam asserted that the evidence shows that it does have significant effects:

> The research reported [in their review article] shows *conclusively* that formative assessment does improve learning. The gains in achievement

appear to be considerable, and...amongst the highest ever reported for international interventions. (Black and Wiliam 1998a, 61) [Emphasis added]

This is a striking claim, because they are suggesting that formative assessment in the classroom actually results in greater success in examinations. Moreover, they point out that 'some, but not all of the studies showed that improved formative assessment helped the (so-called) low attainers more than the rest, and so reduced the spread of attainment while raising it overall' (Black *et al.* 2002, 9). So while pupils of all abilities do better, the apparently weaker pupils make comparatively more progress. This fact in particular suggests that the approach was having a positive effect on motivation; the 'helpless' were becoming 'masters'.

If the approach is apparently so successful, what underpins it? Behind the approach lay a model of pedagogy that drew heavily on the notion of feedback, but how could these be applied in practice? After the original review article, Black and Wiliam identified four key strategies that were important (Black and Wiliam 1998b):

1. Questioning
2. Feedback
3. Sharing criteria
4. Self-assessment (Black *et al.* 2003, 30)

Clearly, the last is the focus of this research, but it should be understood in the wider pedagogical setting.

2.3.1 Questioning

Research on questioning suggests that teachers needed to ask reflective, open-ended questions rather than factual, closed ones: response-seeking not information-seeking (e.g. Stiggins *et al.* 1989). It suggests that teachers should increase the 'wait' time for answers – from the average of less than a second to five seconds. This would result in: longer answers; more thoughtful answers; more pupils answering; pupils being more confident when answering; challenging each other and providing a variety of explanations (e.g. Rowe 1974, Swift and Gooding 1983). The change is from a culture where pupils' oral answers are right or wrong, to one where they are an interesting starting point for learning: the process becomes collaborative. In more recent work this has been re-conceptualised as 'dialogue', as pupils begin to question each other (Alexander 2005, Black and Wiliam 2006).

2.3.2 Feedback

The second strategy also develops a collaborative approach to learning: the use of feedback. This idea is central to Black and Wiliam's theory of assessment for learning. They draw on Sadler's work for a theoretical understanding of this, and he quotes Ramprasad's definition, drawn from engineering:

> Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way (Ramprasad 1983, quoted in Sadler 1989, 120)

Thus, Sadler emphasises, *unless* the information is used to 'alter the gap', it is not genuine feedback: feedback needs to feed-forward. One issue is the form that this information should take; if it is 'too deeply coded (for example, as a summary grade given by the teacher)', it 'may actually be counter-productive' (Sadler 1989, 121). They also draw on Butler's experimental research, for instance, which suggested that the use of summative marks, grades or levels as feedback on work has a negative effect on both pupil motivation and attainment, while giving precise task-related feedback has a positive effect on both. She hypothesised that if pupils are given a 'raw' grade, they will attribute their lack of success to an innate ability (ego-involvement), whereas if they are given clear written feedback, they learn how to improve at the work set (task-involvement). She first studied the effects on the students of either grades alone, or detailed feedback comments alone, and in a second piece of research also studied the effect of *both* comments and grades combined (Butler 1987, Butler 1988, see also Kluger and DeNisi 1996).

For the second research, she conducted an experimental study of 132 Israeli eleven year-olds from four schools, with half of them from the top quartile of their class and the other half from the bottom quartile. Over three tests, she found that grades alone produced no gain in attainment; they had a positive effect on the motivation of the more able but a negative effect on the less able. With

comments alone all the pupils' attainment improved overall, and all were motivated. Comments and grades revealed a *decline* in test scores for attainment, but the more able felt motivated whilst the low achievers were de-motivated. Thus, she not only confirmed that 'task-involving evaluation is more effective than ego-involving evaluation' (Black and Wiliam 1998a, 13), but also showed that the latter can undermine the former if both are used.

The degree of detail in the comments matters. They need be task focused, i.e. what the student has achieved against the success criteria, and what is to be done next, rather than being general praise or ascribing success to an innate quality such as intelligence. Research shows that if the feedback is general, e.g. 'very good', and not task-focused, pupils made less progress than they did with marks alone (Torrance and Pryor 1998, Smith and Gorard 2005) and too much detail was unproductive (Day and Cordon 1993).

2.3.3 Sharing criteria

For feedback to be successful, not only does the teacher need to give clear feedback, but also the pupils need to understand what is required. Pupils need to 'possess a concept of the standard (or goal, or reference level) being aimed for' (Sadler, 1989, 121), i.e. sharing criteria, which is Black and Wiliam's third strategy. The argument is that pupils will only begin to make sense of feedback if they themselves are able to build a picture of what is required. Otherwise, however good the feedback is, they would simply be uncritically following its suggestions without developing a sense of the nature of learning in that subject.

There are two points to make about this strategy. First, there is what Wiliam calls the problem of 'plasticity'; this is 'the impossibility of capturing meaning precisely in criteria' because criteria are written with a particular age group or ability group in mind (Wiliam 1998, 7). This is discussed in my earlier paper in relation to, for instance, the use of a phrase like 'understand key Islamic beliefs'. This criterion could be a description of a learning outcome from Key Stage 2 through to undergraduate level, but in each case the teacher would have a different conception of what it was appropriate 'understanding' at that level, because of their own prior experience. It is also difficult when the criteria are qualitative, or 'fuzzy' (Sadler 1989, 121, see also Radnor 1994 on feedback in art).

Secondly, it should be noted that in the KMOFAP research, this strategy was altered to 'the formative use of summative tests' because in practice, the teachers with whom the researchers worked could not avoid having to use summative processes. They therefore aimed at achieving 'a more positive relationship between the two' (Black *et al.* 2003, 55).

2.3.4 Self-assessment

Finally, we come to the fourth and last strategy: self-assessment. The research evidence suggested that if self-assessment is developed, then it is very effective. As noted earlier, there is no research in religious education on this. There is however research in other areas, both quantitative and qualitative. For example, Fontana and Fernandez describe how Portuguese pupils who were trained in self-assessment in mathematics performed almost *twice* as well as others

(Fontana and Fernandez 1994). Secondly, Fredriksen and White's research into an inquiry-based science curriculum in middle schools from socio-economically deprived areas showed that pupils who were encouraged to discuss and reflect on their learning out-performed the control group. In particular, the previously low attaining pupils made the most progress (Fredriksen and White 1997, in Black and Wiliam 1997, 14-15). Thirdly, McDonald and Boud show that secondary level students in Barbados who were given generic training in self-assessment before taking formal examinations did significantly better than others in *all* their subjects (McDonald and Boud 2003).

Other research suggested the reasons for this. Stephani investigated the correlation between pupil marking and teacher marking in biology. Quantitatively, he found that pupils were more accurate in marking other pupils' work than their own; moreover, qualitatively, he found that *all* students in his study thought that assessing their own or other pupils' work made them think more, and 85% thought that they learnt more (Stephani 1994). Maqsud and Pillai found that pupils who were taught to self-score performed better in tests, and they attributed this to a softening of pupil's attitudes to marking (Maqsud and Pillai 1991).

What is needed for self-assessment to take place? Firstly, Sadler argues, the learner needs to 'possess a concept of the standard (or goal, or reference level) being aimed for' (Sadler 1989, 121). Thus, pupils need to be able to make sense of what is being demanded of them, as described above. Wiliam describes this as 'enculturating learners' into the assessment community (Wiliam

1998). The main problem is what he terms 'plasticity'; pupils need to understand what the expectations of quality are that lie behind the criteria, 'otherwise this is no more helpful than telling an unsuccessful comedian to "be funnier"' (Wiliam 1998, 10). The learners themselves, Wiliam argues, must come to understand the teachers' notions of quality, and therefore teachers will need to develop ways of sharing their notions.

Secondly, the learner also has to 'compare the *actual* (or current) *level of performance* with the standard' (Sadler 1989, 121). Thus, they need to be able to make a judgement about the difference between the requirement and their own efforts. This requires a degree of monitoring of their learning, by comparison. They have to 'picture their own learning in the light of an understanding of what it means to get better' (Black and Wiliam 1998a, 30). This means being able to reflect on their own learning against the concept of the standard that they have come to understand. To achieve this in practice, much of the research on self-assessment included elements of peer-assessment, that is to say that pupils evaluated each other's work. Other people's mistakes are often easier to spot than one's own, and pupils can transfer this critical ability to their own work; it is a 'bridging' skill. For instance, Frederiksen and White's research was based on pupils' discussions of progress, and pupils were commenting on each other's work (Frederiksen and White 1997, in Black and Wiliam 1997). In the classroom, it may be easier to look at particular examples for pupils to comment on each others' work than simply focus on their own work all the time. However, it means that it is not always clear whether the element of individual self-assessment is more or less effective than the element of peer-

assessment.

Lastly, in order for it to be formative, they should 'engage in the appropriate *action* which leads to some closure of the gap' (Sadler 1989, 121); they need to provide themselves with feedforward. There is no point in pupils making the comparison if they simply use it as a measure of their failure or inability to reach, or at least move towards, the standard. They need also to act on their reflections to make improvements. Other definitions of self-assessment do not include this requirement to act (e.g. Boud 1986, 5, quoted in McDonald and Boud 2003, 211), and are arguably weaker without it.

Self-assessment can be seen as the apogee of formative assessment, by bringing together the three other strategies. Moreover, 'self-assessment by the student is not an interesting option or luxury, it has to be seen as essential' (Black and Wiliam 1998a, 54). Firstly, it requires students to ask themselves questions about their own learning, to critique and justify their work. Secondly, it requires them to give themselves feedback, by learning how to identify the remedial action. Thirdly, it requires that they 'understand the standard'. Beyond that, it requires that the pupils be empowered to make their own decisions about how improve; that they are able to control their learning. This, it is argued, will contribute to 'substantive' pupil engagement in education and better motivation, because motivation and engagement in individual lessons will help to consolidate a deeper commitment to education – that "school is for me" (Munns and Woodward 2006, 194). This recalls the earlier discussion about ego-focused and task-focused students; self-assessment helps them become more

task-focused, and arguably gives them the deeper reflective resources to become life-long learners, as McDonald and Boud argue (2003).

2.4 Criticisms of Black and Wiliam

Given the apparent effectiveness of formative assessment, are there any problems with applying it in the classroom? While the broad thrust of Black and Wiliam's original argument is very forceful, especially given the sheer quantity of research that they consider, it would be wrong to suppose that their claims were uncontested, or that their own research did not develop. There are four broad concerns.

2.4.1 Genuine success?

Black and Wiliam's claim in the original article that formative assessment is one of the most successful strategies ever introduced for educational progress is questioned (see section 2.3). Ellwood argues that the original claims should be viewed with 'caution'; she questions the extent of the effects for low attainers, pointing out that 'improvements in scores for low-achieving students are characteristic of many educational innovations and initiatives, not just formative assessment' (Ellwood 2006, 227). She also highlights their own qualifications of their claim:

> We do not claim that formative assessment leads to improved student achievement in *all* cases, with *all* teachers on *all* occasions...Our claim is

> that formative assessment in general is an effective intervention, although we do not estimate the difficulties in translating theory into practice (Black *et al.* 2005, 7 (original emphasis)).

For instance, the KMOFAP researchers found that there were negative effects for four out of the nineteen teachers in their study, and they accepted that the emphasis on quasi-experimental research was both a strength and a potential weakness (Wiliam *et al.* 2004).

Why is this? One reason is that when formative assessment strategies are misunderstood or implemented badly by teachers they can have no impact or a negative impact. Marshall and Drummond (2006) suggest that there is a distinction between the *letter* and the *spirit* of formative assessment, with many teachers using a few techniques, but only about a fifth understanding the deeper processes (see also Smith and Higgins 2007). Similarly, Hattie *et al.* (1996) argue that simply teaching study skills to students without getting them to develop a more general ability to reflect on their learning had no effect. Also, Smith and Gorard's research (2005) shows how the failure to provide *constructive* feedback, which would enable pupils to identify how to improve, to a group of pupils led to worse performances than a control group that simply received marks.

2.4.2 Pupil identity

The failure of some approaches to formative assessment may be the fact that

they require many of the pupils to change their identity as learners, as discussed above. However, they may be unable or unwilling to re-appraise their view of themselves. If whatever teachers have asked them to do in the past has led to their sense of failure, why should they believe that it would be different this time? If they enjoy getting high marks, how will they react when they don't get any? Teachers will have to overcome this hurdle, and this will not be easy. As Perrenoud writes,

> A number of pupils do not aspire to learn as much as possible, but are content to 'get by'...Formative assessment invariably presupposes a shift in this equilibrium point towards more school work, a serious attitude towards learning...Every teacher who wants to practise formative assessment *must reconstruct the teaching contracts so as to counteract the habits acquired by his pupils.* Moreover, some of the children and adolescents with whom he is dealing are imprisoned in the identity of a bad pupil and an opponent. (Perrenoud 1991, 92) (Author's italics)

A change to formative assessment approaches requires both those who were unsuccessful and those who were successful under previous assessment processes to adapt to the new system, and some pupils in both these groups may be unwilling to do so. The change is unlikely to happen overnight, and may well require a whole school strategy.

2.4.3 Other relevant research

Critics questioned whether they had considered *all* the relevant research, which skewed their conclusions. For instance, McDonald and Boud suggest that Black and Wiliam did not refer to research on assessment in higher education, notably on self-assessment (McDonald and Boud 2003, 210). Biggs was concerned with the lack of any research on testing (Biggs 1998), and Perrenoud points out that they only reviewed research in English, arguing that research in French and German would suggest alternative understandings of the process (Perrenoud 1998). Indeed, recent French, Canadian and Swiss research does not distinguish between school and higher education in its conceptualisation of learning and assessment (Allal and Mottier Lopez 2007).

2.4.4 The place of summative assessment

Fourthly, many critics have been concerned with the general conclusions that they draw about the relationship between formative and summative assessment. Biggs argued that as a result of ignoring research on testing, they overplayed the differences. Rather than seeing summative and formative assessment,

> ...as two different trees, I would zoom to a wider angle conceptually. Then in the broad picture of the whole teaching context - incorporating curriculum, teaching itself...and summative assessment - instead of two tree trunks the backside of an elephant appears (Biggs 1998, 108).

Others pointed out the inescapability of teachers having two roles; Sebatane argues that 'multiple assessment techniques should be used, but appropriately' (Sebatane 1998, 125). More recently, Harlen has reframed the relationship as one 'dimension' not a dichotomy, from informal formative assessment, through formal formative to informal summative to formal summative (Harlen 2006, 113-116).

Significantly, the KMOFAP researchers altered the third strand from sharing criteria to the 'formative use of summative assessment', thus zooming to a slightly wider conceptual angle: more of Biggs' elephant's backside. However, the understanding of assessment for learning as *primarily* about this process, seen in recent official documents (e.g.DfES 2003), has been criticised (Daugherty and Ecclestone 2006).

The question is whether pupils can use levels or marks 'ipsatively': 'the principle here is that an individual's performance in a particular domain at a given time should be judged in relation to his or her other performances' (Black 1998, 73). In an ideal world, pupils would be able to appropriate levels and marks in a task-focused way, and would not be tempted to treat them in an ego-focused way. To do this they need to be able to 'decode' them into advice (Sadler 1989, 121).

2.4.5 A narrow view of feedback

Black and Wiliam's conceptualisation of formative assessment, based on Sadler's model of feedback (Sadler 1989), was arguably too narrow. For

instance, Perrenoud drew on Swiss research to suggest that a wider concept of 'the individualised regulation of learning' was more pragmatic (Perrenoud 1998, 85); this would include other regulatory mechanisms, such as 'guidance, control or the adjustment of cognitive, affective and social processes'[1] (Allal 2007, 9). Pillonel and Rouiller have contrasted a narrow view of self-assessment with 'authentic self-assessment'[2]. In the narrow view, pupils are asked '*Judge yourself as I judge you!*[3]' and are trapped into an 'internalised other-assessment'[4] (Pillonel and Rouiller 2001). By contrast, in authentic self-assessment, pupils are asked to '*Judge yourself, however in complete independence of spirit, and not to please or to avoid disagreement with me!*'[5]. Authentic self-assessment arises out of the autonomy of the individual, but it,

> Develops as a result of training that is initiated by the teacher who gives to the pupil a sufficient degree of freedom so that the pupil can take a critical look at his or herself[6] (Pillonel and Rouiller 2001).

Laveault has suggested that there are three pedagogical principles in self-assessment[7]: it must be diversified; it must be differentiated; it must become the

---

[1] Translations NF: le guidage, le contrôle, l'ajustement des activités cognitive, affective et sociales

[2] L'auto-évaluation *authentique*

[3] Jugez-vous comme je vous jugerais!

[4] Une hétéro-évaluation intériorisée

[5] "Jugez-vous cependant en toute indépendance d'esprit et pas pour me faire plaisir ou esquiver tout désaccord avec moi !"

[6] L'auto-évaluation se développe grâce à un apprentissage impulsé par le maître qui accorde à l'élève une part suffisante de liberté afin que ce dernier puisse poser un regard critique sur lui-même.

[7] L'autoévaluation doit être diversifiée...L'autoévaluation doit être différenciée...L'autoévaluation doit faire l'objet d'un apprentissage progressif

object of progressive training (Laveault 2007, 233).

To make practical sense of this, more detailed studies of the micro-level are needed. How would this kind of self-assessment come about in the classroom?

2.5 Formative assessment in the classroom

There have also been demands for a greater understanding of 'the messy point of teachers and pupils responding to each other culturally in relation to classroom discourse and assessment practices' (McFadden and Munns 2002, 364). Indeed, Black and Wiliam had identified the need for research into:

> The nature of the social setting in the classroom, as created by the learning and teaching members and by the constraints of the wider school system as they perceive and evaluate them (Black and Wiliam 1998a, 59).

Although not specifically focussing on assessment, Deakin Crick *et al.* have recently explored the 'ecology of learning' in order to identify all the different factors that are needed to establish 'learner-centred classroom cultures' (Deakin Crick et al. 2007, 267). They include students' perceptions of their own 'learning power', teachers' learner-centred practices, and the emotional literacy of the school culture. Their research across five schools shows student attainment can be affected by the balance of these various factors. Research has also highlighted the importance of whole school approaches in developing the conditions for formative assessment (James *et al.* 2007).

Other studies have considered similar issues for assessment at the classroom level. Cowie investigated assessment for learning in secondary school science in New Zealand, within the socio-cultural dynamic of the classroom (Cowie 2002, 2005). This does not mean wider social or cultural background of the pupils (though that may also be important), but the particular dynamics of the class and lesson. As Cowie argues 'seen in this way, methods of assessment are simply practices which develop patterns of participation that subsequently contribute to pupils' identities as learners and knowers' (Cowie 2005, 140). Her research shows how pupils used a range of criteria to assess their own learning, such as the authority of the teacher, the trustworthiness of their peers and fear of being labelled 'dumb'. This last point highlights the problems of labelling caused by summative assessment practices.

A similar insight is developed by Ecclestone and Pryor, who argue that pupils have an 'assessment career' (Ecclestone and Pryor 2003). This describes an individual pupil's academic, social and emotional experiences of assessment, whether summative or formative, which affects how they will react to further experiences of assessment. They develop Bourdieu's notion of the habitus 'to represent how human knowledge and actions both express and bring about the dynamic of structure in social contexts' (Bourdieu 1977, Ecclestone and Pryor 2003, 473). Pupils 'bring complex dispositions to the field of educational assessment' which 'interact with the structure of the systems themselves and the interpersonal dynamics that arise when teachers and pupils put summative and formative assessment into practice' (Ecclestone and Pryor 2003, 474).

This is an important notion for conceptualising the changing processes that affect pupils' dispositions and motivation. In this respect, pupils' wider experiences of assessment throughout their schooling will need to be considered, in order to understand how they make sense of particular moments in particular subjects.

2.6 Contrasting theories of assessment and learning

The tension between summative and formative assessment, described above, is not just a difference of focus and effect; it is rooted in theories of learning (e.g. James 2006). Roos and Hamilton argue that there are two contrasting psychologies of learning (Roos and Hamilton 2005). On the one hand, summative assessment can be seen as being based on behaviourist theories of learning. It measures pupils' intellectual ability on the basis of their output, as a psychometric process (e.g. Gronlund 1985, Ebel and Frisbie 1991). Seen in this light, assessment is a valuable part of deciding what pupils can do, and is dominated by the test (Gipps 1994).

On the other hand, formative assessment is based on constructivist and social constructivist theories, which focus on developmental aspects of mind. In particular, it is informed by activity theory, notably Vygotsky's notion of the 'zone of proximal development' (Wertsch 1985, Vygotsky 1987). This is the gap between what a pupil can achieve unaided, and what they can achieve with help. He gives the example of two pupils who may currently both be achieving the same level – that of an average eight year old - but one with help might be able

to solve problems characteristic of a nine year old, and the other those of a nineteen year old. Thus, 'the zone of proximal development has more significance for the dynamics of intellectual development and for the success of instruction than does the actual level of development' (Vygotsky 1987, 209). The focus of educational psychology should not be on 'how the child came to be what he is' but on 'how he can become what he not yet is', as Leont'ev, Vygotsky's colleague, observed (Bronfenbrenner 1977, 528). Thus, 'a central feature for the psychological study of instruction is the analysis of the child's potential to raise himself to a higher intellectual level of development through collaboration' (Vygotsky 1987, 210).

Research on formative assessment can be seen as the elaboration of this particular point. How should a teacher collaborate with pupils to maximise the realisation of their potential? How are pupils to be encouraged to do this for themselves? The four key elements of formative assessment, outlined above, can be seen to contribute to this, and we have seen how self-assessment is a vital part of this.

2.7 Framing a general question on self-assessment as assessment for learning

So far the literature review has highlighted a number of issues that a teacher would have to resolve in attempting to introduce self-assessment, but what further issues are raised by the process itself of introducing them? Cowie's research reveals the gap in our understanding in the problems of implementing

assessment for learning in the classroom, and especially self-assessment (Cowie 2002, 2005). How do pupils make sense of what the teacher is trying to instil? What other factors do they judge this against? If assessment for learning is intended to empower students, how and why do they feel empowered? More particularly, what are their 'assessment careers' (Ecclestone and Pryor 2003)? This is to recognise that pupils are not passive recipients of these processes; they may embrace change or they may be cautious of it, or indeed reject it. The general question that emerges is:

- What classroom factors affect self-assessment?

To explore this question further should help unravel some of the reasons how and why self-assessment can succeed or fail.

2.8 Conclusion

This chapter contributes to the research by dealing with a significant issue, both in research and current practice: assessment for learning. The theoretical and research literature on self-assessment has been reviewed, particularly in the context of formative assessment, and the wider horizons of assessment generally. The tensions between summative and formative approaches have been described, in particular the place of high stakes testing in the English and Welsh educational system, and the effect of this on attainment and motivation. The distinction between ego-focused and task-focused pupils has therefore been explored. Arguments in favour of formative assessment were examined, and the

four main elements of formative assessment, according to Black and Wiliam (1998a), were explained: questioning; feedback; sharing criteria; self-assessment. Self-assessment was therefore situated within this wider conceptualisation. Criticisms of Black and Wiliam's ideas were outlined: whether it is a successful as they claim; the problem in challenging pupils' identities as learners; wider research on self-assessment; the place for summative assessment; the limitations in their definition of feedback. Research into the ways that assessment occurs within the classroom has also been explored, particularly pupils' perceptions of how they learn (Cowie 2005), and how they experience assessment (Ecclestone and Pryor 2005), which were also linked to the deeper pedagogical principles that underpin the two approaches to learning.

This has led to the formulation of a general question on self-assessment, which contributes to the originality of the research because it raises the issues relating to assessment for learning, so that they can be related to the issues of pedagogy in religious education. Moreover, it is original in drawing on research in French to comment on developments in the research literature in English (e.g. Allal 2007 and Laveault 2007). From here, it is necessary now to turn to the pedagogy of religious education in order to consider the context within which self assessment will occur.

Chapter Three

Pedagogy, policy and assessment in religious education

Now that the background to self-assessment has been explored, the nature of progress in religious education should be considered, since this is what pupils would have to assess in themselves. The second area for review is therefore the current pedagogy, policy and assessment in religious education in England and Wales. The current national guidelines, entitled 'Religious education: the non-statutory framework' (the 'Framework'), provide a useful starting point for an investigation into current arguments about pedagogy and assessment in the subject (QCA 2004). The Framework is intended to set national standards for religious education by providing local authorities with 'a common starting point' (QCA 2003, 15), but also 'give religious education providers the scope and creativity to enhance teaching and learning' (QCA 2004, 3). It is self-evidently not mandatory; however almost all local authorities rely on it to a greater or lesser extent; two notable exceptions are Hampshire and Birmingham, according to the National Association of Teachers of RE (Blaylock 2007).

The Framework identifies two attainment targets: 'learning about religion' and 'learning from religion'. This distinction sets the scene for a discussion of current requirements across England and Wales; it is however increasingly being drawn upon in international discussions on religious education, after 9/11. For instance, the Organization for Security and Co-operation in Europe (OSCE) has recently published a set of guidelines for its fifty-five member states 'on teaching about

religions and beliefs' (OCSE 2007). It suggests that human-rights based pedagogy for 'teaching about religions' should distinguish between 'learning about religion' and 'learning from religion' (OSCE 2007, 24). Moreover, the Council of Europe is currently reviewing a recommendation to all member states 'on the religious dimension of intercultural education' which promotes the use of this distinction (Council of Europe 2007, explanatory memorandum 34). These documents do not explicitly present them as attainment targets, and indeed do not discuss pupil assessment at all; nevertheless, an understanding of their place in assessment and pedagogy in England and Wales will inform international debates (see Afdal 2007a).

Where do these targets come from, and what happened to them? The Framework is government policy. What follows draws on theories of 'policy scholarship' (Bowe *et al.* 1992), which 'examines the origins and implications of policy change' (Furlong *et al.* 2000, 7); this will provide a critical framework for an understanding of the current policy.

Bowe *et al.* argue that there are three policy contexts that shape policy-making. First, the context of influence: this is where 'interested parties struggle to influence the definition and purposes of education' (Bowe *et al.* 1992, 19). In the case of religious education this would be the political parties and their policy advisory groups, i.e. 'think tanks', the faith communities, and also academics. These influences will possibly conflict with each other.

The second context is that of policy text production. Policy texts *represent* the

policy; they will inevitably reflect the dominant influences, but they are also formal, official documents that need to bear the weight of political or legal scrutiny, and relate to other policy documents. They are 'the outcome of struggle and compromise', and while they attempt to be coherent, there may be contradictions and ambiguities in them (Bowe *et al.* 1992, 21).

The coherence or incoherence of the text has important consequences in the third context, that of practice: 'policy is not simply received and implemented...it is subject to interpretation and then 'recreated', because 'practitioners do not confront policy texts as naïve readers; they come with histories, with experience, with values and purposes of their own' (Bowe *et al.* 1992, 22). Schools and teachers will interpret the policy in their context, and implement it accordingly.

Bowe *et al.* draw on the work of Barthes, to distinguish between 'readerly' and 'writerly' texts (Barthes 1970). A 'readerly' text gives the reader little freedom other than to accept (or reject) the meaning in the text. A 'writerly' text by contrast invites the reader to join in, to 'co-write' the meaning of the text; Barthes' distinction is between the works of nineteenth century novelists, e.g. Balzac (readerly), and avant-garde novelists, e.g. Robbe-Grillet (writerly): *Oliver Twist* v *Finnegans Wake*. Bowe *et al.* apply this distinction to policy texts. Some texts allow for a 'writerly' approach, so that teachers are entitled to interpret them freely; other texts are 'readerly', and their meanings accepted. Hargreaves contrasts this as a 'tension between *mandates* and *menus*', the one a command, the other a choice (Hargreaves 1997, 345, author's emphasis). With readerly texts, teachers can find themselves determined by external forces that

metamorphose them into something they do not want to be, which Parkison strikingly compares to Samsa's fate in Kafka's story, *The Metamorphosis* (Parkison 2008). The process does not end there. In a circular way, teachers' implementation of the policies is scrutinised by agents in the context of influence, and new policy is produced, by agents in the context of policy text production. Moreover, the relationships between these contexts are fluid and can operate in a variety of directions, and teacher's voices can also be one voice in the context of influence.

This threefold distinction is further complicated by the fact that there is a middle tier of local Standing Advisory Councils on Religious Education between central government and teachers ('SACREs'), which have the legal responsibility to produce the local agreed syllabus. These SACREs are themselves a context of practice for the implementation of the text and they will each read it in different ways. However there is also a context of influence, as different parties seek to dictate the shape and form of the locally agreed syllabus in the light of the Framework, and moreover the text of this local syllabus will itself need to be produced. Thus the Framework will not be directly translated into practice but will be filtered through this process. Further, teachers can read the Framework itself, to aid their interpretation of their agreed syllabus.

The distinction between 'learning about religion' and 'learning from religion' is fundamental to current policy, and its history can be told in the light of these contexts. Copley has described the broader sweep of the history of religious education, from 1944 to 1995 (Copley 1997), and Robson focused on the political

and legal changes from 1985 to 1995 (Robson 1996, see also Jackson 2000a). The focus here is on the textual history of this distinction.

The distinction was originally made in 1981 by Michael Grimmitt in an article entitled 'When is "Commitment" a problem in religious education?' which discussed the inter-play of personal and professional values for teachers of religious education (Grimmitt 1981). Grimmitt then developed this distinction pedagogically in various works over six years (e.g. Grimmitt 1982), and its most developed exposition is in his book *Religious Education and Human Development* (Grimmitt 1987). After that, the distinction developed a life of its own, of which Grimmitt did not always approve (Grimmitt 2000).

Each of the two current attainment targets is considered separately, and then how these are situated in a wider understanding of the purpose and goal of religious education is discussed.

3.1 'Learning about religion'

First, what is 'learning about religion'? Grimmitt originally suggested that:

> The methodology of Religious Studies provides the basis for *learning about* religion and the type of interaction here is precisely that endorsed by phenomenology (i.e. studying religions objectively 'in a descriptive, critical and experiential manner') (Grimmitt 1981, 49).

The two important elements of this are that it is based in phenomenology and that he suggests that it is the study of 'religions'. Later, he defined it more fully:

> When I speak of pupils *learning about* religion, I am referring to what the pupils learn about the beliefs, teaching and practices of the great religious traditions of the world. I am also referring to what pupils learn about the nature and demands of ultimate questions...about the normative views of the human condition, about the discernment and interpretation of *Core Values*...This type of learning can be said to be an impersonal or public mode of understanding' (Grimmitt, 1987, 225) (author's italics).

Thus learning about religion meant the study of world religions, and the drawing out of general issues of the human condition and central values common to all of them. This drew on phenomenological approaches which considered that study of different religions led one to be able to understand 'religion' as a whole, and religion as a whole would be manifested in particular religions (see Grimmitt 1991). It also depends on a notion of 'Core Values'; these are universal and implicit within the 'givens of the human situation and act as kinds of "value-imperatives"' (Grimmitt 1987, 121). He maintained that 'it is generally recognised that the identification of "key" concepts of religion and their specific interpretation within particular religions is an essential pre-requisite to curriculum decision-making in religious education' (Grimmitt 1987, 129).

He called these key concepts 'substantive religious categories'. Even though he recognised that his own list is drawn from a European perspective, he

nevertheless considers that it is possible and desirable to compile it. Grimmitt was not arguing for a kind of Christian theological pluralism - that all religions were simply different paths to the same God; unlike Hick, who argued that Christians needed to 'see both our own and other great world religions as revolving around the same divine reality' (Hick 1983, 82-83). Further, Grimmitt was also critical of experiential approaches to religious education developed by David Hay and his colleagues. These approaches were based on the work of the Religious Experience Research Unit, which showed that many people in Britain claimed to have had a religious experience, in the sense of an awareness of a transcendent power (Hardy 1979, Hay 1982). Hay argued that *everyone* had an innate religious sense, which particular religions both awakened and formalised. The role of religious education was to raise pupils' awareness of their own inner spirituality as a vehicle for the exploration of religions (Hay 2000); Grimmitt was cautious of this claim (Grimmitt 1987, 180-193). Nevertheless, one can see how both arise out of similar essentialist views of religion.

3.1.1 Learning about religion - The Model Syllabuses

'Learning about religion' entered policy texts as attainment targets in 1994, in the government's model syllabuses (School Curriculum and Assessment Authority ('SCAA') 1994) (the 'Model Syllabuses'). This was because in the early 1990s the learning from/learning about polarity was taken up by exam boards; they were used in the general criteria for GCSE (Attfield 1996). At the same time assessment came to the fore in religious education after the Education Reform Act 1988, when the National Curriculum in other subjects included assessment

scales (Religious Education Council 1991).

Two major projects developing the role of assessment in religious education were developed: the FARE project and the Westhill project; these were designed to *influence* practice. The FARE project involved a hundred teachers across eight local authorities. It identified two attainment targets: 'reflection on meaning' and 'knowledge and understanding of religion' (Copley *et al.* 1991). The Westhill project involved twenty-nine local authorities. It originally identified *ten* attainment targets (Westhill College 1989), but this number was impractical (Blaylock 2000, 46) and was condensed to three: 'understanding religious belief and practice', 'reflecting, responding and expressing' and 'awareness of life experiences and questions they raise' (Westhill College 1991). Neither used Grimmitt's terminology. Like Grimmitt, the FARE project identified two targets, but Grimmitt's pedagogy more strongly influenced the Westhill project (Rudge 2000b).

What happened when the distinction entered the policy texts? At this point, the context of text production becomes important. In discussing this process, Barbara Wintersgill claimed that this type of document,

> ...must by definition present a compromise position because they have to be acceptable to the whole community – particularly the faith communities. I say 'particularly' because the government will not support any national RE document that is not acceptable to them – they are political players. Frankly the government doesn't [care] what AREAIC,

> teacher trainers or even less academics think of these documents! This may sound cynical but I have always believed that within this nonsense, if only the RE world was more politically astute, we can bring about a compromise. Researchers and academics take our ideas forward and these do inform national documents because the officers who draft them DO care what AREAIC etc think. (Wintersgill 2007)

This shows how in her experience, the different areas of influence can affect different stages of the text production process. The views of particular members of the faith communities were influential on the direction of government policy, notably Baroness Cox (Cox 1988, Palmer 1993); however, the academic debates influenced the document writers, who had an eye to the context of practice.

It would be wrong however to see Thatcherite government influences on education at this time as univocal. Furlong *et al.* suggest that on the one hand, there were neo-liberal perspectives which had led to the introduction of free market principles in education, including the use of assessment results for drawing up league tables (Furlong *et al.* 2000, 9-12, see Torrance 1997). On the other hand there were neo-conservatives, notably the Hillgate group, who wished to preserve a 'refined cultural heritage' (Furlong *et al.* 2000, 11). Goodson argues that these can be seen as twin responses to globalization (Goodson 1990, in Hargreaves 1997). Religious education was 'caught up in the unfinished debate about British national identity' (Robson 1996, 16).

The Model Syllabuses were influenced by both tendencies. They can be seen as

part of a neo-liberal rationalization of delivery in schools to create an educational market, but were also part of a neo-conservative reaction based on a perception that religious education was becoming 'multi-faith relativism' and that Christianity, which deserved a special place, was being ignored (e.g. Burn and Hart 1988, to which Baroness Cox wrote the Foreword). These two tendencies are not necessarily coherent; Burn and Hart are very critical of the neo-liberal tendencies.

In terms of the detail of the Model Syllabuses, this aspect of the subject was labelled 'learning about religions' (in the plural), and it was described in terms of the study of individual religions, including comparing religions. Thus, it should include the ability to:

> Identify, name, describe and give accounts in order to build a coherent picture of each religion;
> Explain the meaning of religious language, stories and symbols;
> Explain similarities and *differences within*, and between, religions. (SCAA 1994, 5) (Emphasis added)

This new approach found favour with some in the academic community because it focused on the study of discrete religious traditions and not on the 'essence' of religion. For instance, Jackson had criticised the unquestioning use of the word 'religion' in two ways (Jackson 1993, Jackson 1997, 49-71). The first is whether 'religion' can be conceived of as an entirety. He argues that the development of notion of 'religion' as a generic category in the nineteen-century was an idealist,

particularly Hegelian, notion; there was something that could be called 'religion' that was greater than, but manifested in, particular religions. Against this, Jackson argues 'for a general loosening of established categories....which posit universal 'essences" (Jackson 1997, 69).

His second point was whether it is appropriate to use the term 'religion' even to describe the individual systems of beliefs, values and culture such as Christianity, Islam or Hinduism. Arising out of his research into British Hindus, Jackson argued that even the notion of an individual 'religion' is suspect. This is because it tends to both overlook internal diversity within an apparently unified whole and it tends to assume that all so-called religions are the same sort of entities. While he does not think that the word should be abandoned altogether, he is concerned that such approaches 'entrap insiders within schematic formulations of key beliefs and concepts' (Jackson 1997, 69).

The Model Syllabuses reflected this concern. There was no suggestion that an abstract sense of 'religion' should be developed. Further, although the first element in this assumes that one can build up a coherent picture of each religion, the sense of internal diversity was at least recognised in the third element. Neo-conservative influences at the time were less unhappy with a plural traditions approach because they were wary of an incipient theological pluralism in the essentialist view (Blaylock 2007, Wintersgill 2007). However, if the Model Syllabuses reflected this concern, they also created some ambiguity, because 'religion' was undefined (Everington 2000). Was 'religion' simply the sum of the six religions identified or something wider? (Hull 1995); thus despite the neo-

conservative influences, some 'Christian traditionalists' were unhappy with the final text (Hull 1994).

What happened in the context of practice? There does not appear to be *direct* research on how teachers implemented these. There was textual ambiguity, and thus a more writerly text, which teachers were likely to interpret widely. Bearing in mind Sikes and Everington's research (2000) which highlights the different personal reasons for becoming a teacher of religious education, it is likely that there would be a range of responses. For one group of teacher trainers, the text appeared too mandatory, and they produced their own third 'dimension', concerned with human experience (Baumfield *et al.* 1993, 1994), though Wintersgill argued that this issue was to be addressed by teachers in the classroom (Wintersgill 1994). Blaylock argued that teachers should develop 'authentic' assessment, linked to classroom practice, which can be seen as arguing for a writerly approach (Blaylock 2000). The SACREs would limit their effect because they would be filtered through the local syllabuses; South Gloucestershire for instance modelled their 2000 syllabus on the FARE project (South Gloucestershire County Council 2000).

Other research addressing differences in classroom practice is also relevant; it points to differences in classroom practices. Teachers would use a variety of *methods* to deliver different aims in the Model Syllabuses (Astley *et al.* 1997). Furthermore, that the Government was aware of differences in practice was shown by the fact that it made detailed knowledge of the Model Syllabuses compulsory for trainee teachers in religious education (Grimmitt 2000b, 14), and

also by the publication of exemplification material (QCA 1998). Finally, although not directly concerned with the situation in England and Wales, German research suggests that different styles of delivery by teachers will also affect pupils' 'inter-religious learning' (Gennerich 2007).

3.1.2 Learning about religion - Non-statutory Guidance

The next stage in the attainment target's history was the publication in 2000 of a guidance document to assessment in religious education by SCAA's replacement, the Qualifications and Curriculum Authority (QCA 2000) (the 'Guidance'). This was 'to help improve consistency and effectiveness of assessment' (QCA 2000, 4). It did not change the generic descriptors from 1994, but did set out an eight-level scale against which pupils up to the end of Key Stage 3 (aged fourteen) could be summatively assessed; it was based on AREIAC's model (AREIAC 1988), and QCA's own exemplification material (QCA 1998). It followed the pattern for other subjects in the National Curriculum, which was also revised that year, including specifying median attainment levels for the end of each Key Stage (QCA 2000). The criteria were tightly specified for each of the eight levels (plus one for 'exceptional performance'); indeed it sub-divided them for each of the three sub-sections for each attainment target: fifty-four descriptors in total. This tends towards 'hyper-specification' (Wiliam 1998, 7) and can be seen as a readerly text.

The first attainment target was divided into three sub-elements:

> Knowledge and understanding of religious beliefs and teachings;
>
> Knowledge and understanding of religious practices and lifestyles;
>
> Knowledge and understanding of ways of expressing meaning. (QCA 2000, 4)

These however do not match the original three sub-elements from the Model Syllabuses (see above). For instance the original third element was 'to explain similarities and differences between and within religions' (SCAA 1994, 6); this was reduced to one part of level five for the first new sub-element ('explain how some beliefs and teachings are shared by different religions' (QCA 2000, 9)).

Moreover, although the scale could be seen as being an important part of 'raising standards', there were three issues with it. Firstly, it seemed to suffer from the problem of hyper-specification that had bedevilled the 1988 National Curriculum, which was 'an over elaborate system which...serves to fragment teaching and learning' (Dearing 1994, 61, in Blaylock 2000, 46).

Secondly, although the assessment criteria were now specified nationally, the curriculum was not, being set locally; thus there was still a secondary level of text production. Further, pedagogy and assessment must be aligned for assessment to be valid (James 2006, 47), but they were not because it was 'difficult to make a general statement' of what pupils might be expected to know (Kay 2005, 44). Local authorities tried to resolve this (e.g. Devon Local Education Authority undated, Wakefield Local Education Authority 2001), and created a new level of interpretation.

Thirdly, the descriptors for particular levels were not sequential. For example, level five required pupils to 'explain how some principal beliefs...are shared by different religions' and level six required them to 'explain how the principal beliefs and teachings...vary among different groupings, denominations and traditions'. It is however not clear why identifying similarities between different religions is a more fundamental process than identifying differences within one religion. The pupils' outcome could be dependent on the nature of task set, not the quality of what they produced. One problem that John Keast encountered when drafting it was that certain command words, e.g. evaluate, could only be utilised by QCA at a certain level in any National Curriculum subject, possibly influenced by Bloom's taxonomy (Bloom 1956, Keast 2003b). These served to skew the pedagogical aims, and the text was therefore more open to more writerly than readerly approaches, as much depended on how the teacher interpreted the levels in the context of their teaching.

Thirdly, there were also wider pedagogical issues being discussed. Wright argued for a theological sense to 'religion', stating that 'the vertical dimension of religion – the fact that in phenomenological terms, the distinctively "religious" dimension of any given phenomenon lies in its orientation or intentionality towards transcendence – cannot be dismissed' (Wright 2004, 197). He considers that it is not 'meaningless' to talk about religion in a general sense, and is particularly concerned that a merely descriptive, phenomenological approach will not do justice to the truth claims that pupils should be tackling. This approach aims at a quasi-objective description of the features of religion and not

the 'disputed questions of faith, value and commitment' (Wright, 2004, 186).

By contrast, other writers thought that the notion of religion should be abandoned altogether. For instance, Clive Erricker argued for a 'post-modern spiritual pedagogy', which meant not only abandoning the reified concept of religion, but also recognising the inadequacy of focusing on material from religious traditions:

> Children's spirituality cannot be expressed and reflected on by the children themselves as long as the subject enquired into is defined by the concept 'religion' (Erricker and Erricker, 2000, 26).

This is a radical solution. It is also worth highlighting the issue of spirituality, as this is also seen by many as an important element in the pedagogy of religious education. In this research it is not possible to delve deeper, as it raises another body of literature.

The Guidance had provided a national set of assessment criteria. It had not explicitly tackled the pedagogical issues that the Model Syllabuses had raised, but it had provided an implicit re-interpretation of them. Assessment and pedagogy still needed to be combined and aligned.

3.1.3 Learning about religion - the Framework

In 2004, the Framework was published, but did it resolve these problems? It defines learning about religion (in the singular) as:

> Enquiry into, and investigation of, the nature of religion, its beliefs, teachings and ways of life, sources, practices and forms of expression. It includes the skills of interpretation, analysis and explanation. Pupils learn to communicate their knowledge and understanding using specialist vocabulary. It also includes identifying and developing an understanding of ultimate questions and ethical issues. In the national framework, learning about religion covers pupils' knowledge and understanding of individual religions and how they relate to each other as well as the study of the nature and characteristics of religion (QCA 2004, 11).

It avoids the three sub-elements structure, preferring to set out a range of outcomes. Overall, there are two problematic areas: the issues of hyper-specification and plasticity, and the debate about 'religion'.

First, the Framework is structurally less tightly specified than the Guidance because it does not identify criteria for the six sub-elements individually (QCA 2004, 36). It still separates the two targets, with the eight levels plus exceptional performance for each one (eighteen in total). Some academics were critical of this separation, e.g. Jackson, arguing that 'the perpetuation of the separate attainment targets – learning about religion and learning from religion – invites teaching to one or the other, rather than integrating different elements of learning' (Jackson 2004a, 220).

Secondly, one can see this as opening the way for a more essentialist concept of 'religion'. In this respect it is closer to Grimmitt, than the Model Syllabuses.

There had been a major change in the context of influence, from a Conservative to a Labour government. While New Labour maintained many of the neo-liberal elements of education (Grimmitt 2000b, 10), the voice of theological neo-conservatism was missing, and this may have meant that the pressure not to include the concept was absent. However, it does not ground this in the 'Core Values' that are part of Grimmitt's 1987 version, or identify any 'substantive religious categories'.

Critically, 'religion' remains undefined (Maybury and Teece 2005). Thus, a theologically conservative criticism is that the framework is ambivalent, requiring the study of 'the nature of religion', and study of 'individual religions', without specifying their inter-relationship (Felderhof 2004, 246). This is particularly ironic as the Framework also controversially allows for secular world views to be studied in the subject (QCA 2004, 25-29). At the same time, Jackson was concerned with the lack of any sense of the dynamic changing cultural forms of religious traditions, considering it still too static (Jackson 2004, 220).

In the text, there is a tension between readerly and writerly elements: assessment levels are very prescribed, but 'religion' is not. Compromises at the level of text production in the light of a variety of influences can create openings for a variety of interpretations in the context of practice. One might assume that those involved in text production aimed at readerly texts, but this is not necessarily the case. Wintersgill commented that,

> National documents...are broad and flexible enough to allow for

> interpretation – indeed the best RE I have seen has been in schools where the department has not followed unthinkingly a text book or syllabus but has brought its own ideas to bear (Wintersgill 2007).

Further evidence of the diversity of readings of the text can be seen in Hayward's recent survey of the teaching of Christianity at Key Stage 3 (Hayward 2007), in which teachers' ideas for 'learning about Christianity' could include a range of concerns, such as its global presence, as a way of life, or Jesus' teachings.

3.2 'Learning from religion'

How does the second attainment target fare? The key dates and publications are obviously the same, but the target itself has been even more widely interpreted. Grimmitt originally wrote:

> The other type of interaction envisaged - *learning from* religion - encompasses and builds on the type of understanding promoted by Religious Studies. More accurately, it is concerned with the *evaluation* of one's understanding of religion in personal terms and the *evaluation* of one's understanding of self in religious terms. (Grimmitt 1981, 49, (author's italics))

This is because:

> The value of the study of religion from an educational point of view is to

> be located in the insights it provides into one of man's most characteristic activities - his search for meaning and his involvement in meaning making and truth-making. These insights are transferable to the pupils' existential situation through 'the process of religious education' (Grimmitt 1981, 49).

The use of the singular word is central to Grimmitt's understanding of the interrelationship between the two elements. Both 'Religion' and the pupils' lives are existentially rooted in the human situation, and 'Religion' can therefore offer insights to pupils.

By his later book this had been developed:

> When I speak of pupils *learning from* religion, I am referring to what pupils learn from their studies in religion about themselves – about discerning ultimate questions and signals of transcendence in their own experience and considering how they might respond to them, about discerning *Core Values* and learning to interpret them, about recognising the shaping influences of their own beliefs and values on their development as persons, about the unavoidability of their holding beliefs and values and making faith responses, about the possibility of their being able to discern a spiritual dimension in their own experience, about the need for them to take responsibility for their own decision-making, especially in matters of personal beliefs and conduct, and so on (Grimmitt 1987, 225 (author's italics)).

This is more explicit as to how the pupils will achieve these insights. It also links the process in the 'Core Values', which are, for Grimmitt, existentially shared.

One part of his definition is about evaluation. He remarks that pupils should,

> ...evaluate the claims of religion and a religious interpretation of life by engaging in an open and critical exploration of the interface between what they perceive to be the central teachings of religions and their own questions, feelings, experiences and ideas about life (Grimmitt 1987, 203).

He goes on to distinguish between: 'Impersonal Evaluation', as the critical consideration of the material studied such as truth claims, beliefs and practices; 'Personal Evaluation', as the critical consideration of one's own beliefs in the light of these. He suggests that certain kinds of questions will aid the promotion of such 'self-knowledge': those of self-identity, e.g. 'who am I?'; self-acceptance, e.g. 'how do others see me?'; self-illumination, e.g. 'why am I like I am?'; self-ideal, e.g. 'what is the right thing to for me to do?'; self-adjustment, e.g. 'should I change?'; self-evaluation, e.g. 'how do I feel about myself?', and then 'who am I?' again (Grimmitt 1987, 227-228).

This reflective, evaluative dimension was echoed in the two other models of assessment. In the FARE project, this aspect was termed 'reflection on meaning', and divided into three elements; awareness of mystery, questions of meaning, and values and commitments (Copley *et al.* 1991, 35). In the Westhill

project, it included elements of 'reflecting responding and expressing' and 'awareness of life experiences and the questions they raise' (Westhill College 1991).

3.2.1 Learning from religion - the Model Syllabuses

The next major step was its inclusion in the Model Syllabuses, in 1994. SCAA described it as:

> Learning from religion
>
> This includes the ability to:
>
> • Give an informed and considered response to religious and moral issues;
>
> • Reflect on what might be learnt from religions in the light of one's own beliefs and experience;
>
> • Identify and respond to questions of meaning within religions (SCAA 1994, 5)

The first striking point 'religion' that it is in the singular; thus the two attainment targets are misaligned: the first with 'religions', second with 'religion'. This however does not seem to reflect a shift in pedagogical position, since the second and third sub-elements of the second refer to 'religions', and the first sub-element is simply neutral in this respect. The second point is that this is the language of compromise - it could after all *include* other processes, such as Grimmitt's model. According to Wintersgill, this attainment target was nearly

scrapped just before publication of the document because the neo-conservative government-appointed representatives were suspicious of this learning outcome, seeing it as a tendency towards 'mishmash' (Wintersgill 2007). As written, it is a narrower conception than either of Grimmitt's descriptions. He talks of a wider, more existential process, such as 'discerning...signals of transcendence', 'recognising the shaping influences of their beliefs' on personal development, or 'being able to discern a spiritual dimension'.

This wider sense of learning from religion was however not reflected in the detailed descriptors for each key stage. They subdivided it differently, into two sub-elements; pupils should have opportunities for 'enhancing their own spiritual and moral development' and 'developing positive attitudes towards other people and their right to hold different beliefs' (SCAA 1994, 13, 25, 37, 49). There was therefore a major inconsistency in the document as to what this target meant, which opened the way to writerly interpretation.

Even in its widest sense, not everyone approved of it. Attfield's theologically conservative criticism was that while Grimmitt's first attainment target was 'unproblematic', the second could result in 'a personal appropriation...of a mixture of elements from various faiths' and that 'it is hazardous to claim that this syncretistic form of personal growth is desirable' (Attfield 1996, 81). He argued instead for a faith development model within one tradition, though he accepted that pupils might reject this (see also Ellis 1997, Attfield 1997). Unfortunately, there does not appear to be any empirical evidence as to how teachers interpreted it in the context of practice.

3.2.2 Learning from religion - Non-statutory Guidance

In 2000, the QCA produced the Guidance, which set out the assessment levels for the SCAA attainment targets, as described above. QCA felt it necessary to devote twenty out of thirty-five pages to 'learning from religion' because 'monitoring of agreed syllabuses by QCA and inspection of RE by OFSTED both indicate that much of pupils' learning about religions is not accompanied by planned learning from these religions' (QCA 2000, 16). This suggests that the text was far from readerly, except that teachers were not so much interpreting it for themselves as abandoning it altogether as an attainment target.

First, did the eight-level scale resolve the issues? In order to specify the criteria for eight levels, QCA introduced three new sub-elements within the attainment target:

> Skill of asking and responding to questions of identity and experience;
> Skill of asking and responding to questions of meaning and purpose;
> Skill of asking and responding to questions of values and commitments.
> (QCA 2000, 4)

These introduce the language of 'skills' (for a critique, see Kay 2005, 44-46). They also restructure the 1994 generic threefold sub-elements, but not the two-fold division given at each key stage. For instance, the first sub-element in 1994 ('give an informed and considered response to religious and moral issues'), seems to be rephrased as the third new sub-element, but the idea of 'response'

has been applied across all three sub-elements, and the skill of 'asking' questions has been added from the original third sub-element ('identify and respond to...'). As far as the level descriptors themselves are concerned, the problem of the 'plasticity' of criteria is particularly acute because the descriptors can only describe assessment norms, which were still unclear (Fancourt 2005), and Kay has pointed out the lack of obvious progression in them (Kay 2005, 42-43).

What of the specific guidance on learning from religion? This includes a number of suggestions. Some appear to resemble 'learning about religion', e.g. 'understand how believers in different religious traditions may interact with each other', and others echo the Westhill project's threefold structure, e.g. 'pupils should be able to make clear links between *common human experience* and what religious people believe and do' (QCA 2000, 16 (emphasis added)). Good practice is said to include being 'inextricably linked to ...learning from religions' and 'about concepts in religions', and moreover 'about developing skills, e.g. the skill of living in a plural society' and open-mindedness, whereas bad practice is 'free of religious content...about promoting a religious lifestyle...an invasion of pupils' privacy' (QCA 2000, 18). As late as 2003, I heard a teacher trainer describe it as the promotion of Christianity. It is also said to be 'essential to the spiritual, moral, social and cultural development of pupils', which echoed the end of Key Stage language in the Model Syllabuses (QCA 2000, 23).

What emerges from this document overall is a tension between the hyper-specification of the scale and the more general pedagogical goals of the subject.

The guidance is intended to provide detailed sequential descriptors of assessment criteria but also to hold true to a wider rationale for religious education. This is a reflection of a tension between the use of assessment in neo-liberal vision of measurable standards and a more personal approach to assessment in the subject. It reaches its nadir in the section on 'assessing learning from religion'. Thus, while 'not all aspects of learning from religion are assessable or measurable' and 'the personal views, attitudes or emotions of pupils are not assessable and it would be intrusive to do so', nevertheless 'how well pupils demonstrate they are developing the skills of responding and evaluation is assessable' (QCA 2000, 23). Finally, a wider conception of assessment is also expounded:

> Not all assessment in RE consists of tests or external judgement. Self-assessment is particularly useful in some aspects of learning from religion (QCA 2000, 23)

This tantalising new view of assessment is not developed further: which aspects is it useful for? Is this the same as or different from the use by pupils of the scale? How is this process envisaged? This glimpse of other possibilities opens up a writerly approach; while it sets limits on the role and place of some types of assessment, it leaves these open to interpretation by teachers.

How did Grimmitt view this? In 2000, the same year as QCA's guidance, he reviewed the state of the subject, and was not pleased with the way that his vision had been reduced. He does not refer to the QCA's guidance, but was

possibly aware of the AREIAC version:

> Where teachers attempt to encourage pupils to *learn from religions* [*sic*] it falls far short of the pedagogical strategy which the original concept involved, namely that pupils should evaluate their understanding of religion in personal terms and evaluate their understanding of self in religious terms (Grimmitt 2000, 15).

It was not however simply that his idea had been watered down that concerned him, it was the wider policy agenda within which this had taken place. As quoted earlier (section 2.1), he lambasted 'a technicist and standards-related political ideology of education'. His two aspects of learning about religion and learning from religion had now become 'learning outcomes' and 'attainment targets' in a culture of 'standards' that could be assessed and measured. He laments this shift:

> The traditional language of *learning* and *teaching* was absorbed and lost within a framework dominated by the notion that by clearly stating the *intended outcomes* of any educational or training programme this would enable objective judgements to be made about the success or otherwise of pupils and students, and, by implication, the success of schools and the teachers which they employed (Grimmitt, 2000, 9).

The key point here is that a newly imposed educational policy appropriated previous approaches to the subject. The context of text production drew on

academic influences to flesh out the demands of government, as Wintersgill suggests; in this case they used Grimmitt's terms but transposed them into a different, official paradigm.

As far as the academic influences are concerned, there are of course other pedagogical models of how these reflective processes might work. Thus, Jackson is broadly sympathetic to Grimmitt, considering that their views have 'more similarities than differences' (Jackson, 2004b, 135). In particular, he has drawn on a hermeneutic model of learning that seeks to develop a dialogue between the pupil and the material studied. For 'learning from religion', Jackson's use of ethnography as a model for his pedagogy means that he uses the sociological term 'reflexivity' to describe this, defined as 'the relationship between the experience of the students and the experience of insiders whose way of life they are studying' (Jackson 2000b, 134). He identifies three aspects of this:

1. Learners re-assessing their understanding of their *own* way of life (being 'edified' through reflecting on another's way of life)
2. Making a constructive critique of the material studied at a distance
3. Developing a running critique of the interpretive process – being methodologically self-critical (Jackson 2004, 88).

Thus, what Grimmitt calls impersonal evaluation, Jackson labels as 'constructive criticism' of the material studied, and what Grimmitt calls personal evaluation, Jackson talks of as 'edification'. Incidentally, Jackson's definition echoes the

long-standing idea of self–assessment (in point one) as a wider existential process, which will be discussed later. He also emphasises how it is inseparable from the process of 'grasping another's way of life' (Jackson, 2000b, 135).

While Jackson is supportive of Grimmitt but critical of the Framework, Wright is critical of all three. He considers that 'learning from religion' is based on a false notion that 'ultimate truth is to be found within, in that private inner space where we encounter our primal desires' (Wright 2004, 187). However, Wright *is* also concerned that the pupils' own perspectives are taken into account in a 'critical' religious education.

> It is vital to effective education that children are given the freedom and encouragement to recognise and articulate their emergent religious beliefs and attitudes without constraint or manipulation, and that the horizons of the pupils - whatever their material content – become a conscious and integral part of the learning process (Wright 2000, 178).

Moreover, like Jackson, he considers that this needs to be integrated into teaching rather than being a separate process from the investigation of the material studied.

> Learning takes place...when the horizons of the pupil and of religion are brought face-to-face...Critical education demands the cultivation of intelligent conversation between the two horizons, in which the religious horizons of the child encounters a range of religious perspectives and

options, and uses this encounter as a means of further clarifying, enriching and developing the child's religious beliefs (Wright 2000, 179).

As far as the academic influences are concerned, there are three issues that arise leading up to the Framework. Firstly, there is agreement that the development of pupils' own perspectives on religion(s) is a vital dimension to the subject. Secondly, the relationship between learning from and learning about is both contested and, in government documents, confusing. Thirdly, there is a tension between the demands of standards-driven hyper-specification and the subject's wider goals.

3.2.3 Learning from religion - the Framework

Does the Framework resolve these issues? Learning from religion is:

> Concerned with developing pupils' reflection on and response to their own and others' experiences in the light of their learning about religion. It develops pupils' skills of application, interpretation and evaluation of what they learn about religion. Pupils learn to develop and communicate their own ideas, particularly in relation to questions of identity and belonging, meaning, purpose and truth, and values and commitments. (QCA 2004, 11, see also 34)

First, following on from the problems with 'learning about religion', it is argued that the lack of a definition of religion leaves a conceptual whole at the centre of

this attainment target, and therefore the subject (Maybury and Teece 2006). Do pupils learn from religion in the same way that they learn from religions? And will they learn from secular philosophies in the same way? (Watson 2007).

Secondly, the Framework can be seen as being more coherent than the Guidance, in that it uses 'religion' consistently, but it still maintains the distinction. Some object to this, e.g. Jackson (2004a), and Copley recently argued for a single target 'engaging with religions and other life stances' (Copley 2005, 128), though it is not clear how this is consistent with his earlier advocacy of a twin target approach in the FARE project (Copley *et al.* 1991). OFSTED has also recently questioned its value (OFSTED 2007, 39).

On the third point, the Framework maintains a more summative paradigm of assessment, because it closely follows the revised National Curriculum; at the level of text production, coherence between different policy texts seems to have taken precedence over concerns with the context of practice. This is shown in the omissions, e.g. there is no advice on different types of assessment, as in the Guidance. What it does suggest is that 'assessment needs to take place in relation to both attainment targets', and that teachers should do this (QCA 2004, 35). Further, it simply adds:

> It is important to note that not all aspects of religious education can be assessed. For example, pupils may express personal views and ideas that, although integral to teaching and learning, would not be appropriate for formal assessment. (QCA 2004, 35)

Thus the Framework merely limits the use of 'formal' assessment, whereas in the Guidance the use of self-assessment was encouraged. Moreover, it assumes that its use in relation to 'learning about religion' is unproblematic, but the earlier discussion on the problems with summative assessment in the classroom might suggest otherwise. This lack is particularly surprising because the government had also been urging the greater use of assessment for learning (Department for Education and Skills 2003). It is particularly ironic because the Foreword, by the Secretary of State for Education at the time and the chief executive of QCA, states that 'Good-quality religious education can transform pupils' *assessment of themselves* and others, and their understanding of the wider position of the world in which we live' (QCA 2004, 3) (emphasis added). Thus, much of what is demanded of the subject arises in relation to the wider outcomes of religious education, and these deserve special discussion before considering the context of practice.

3.3 Attitudes and values in religious education

In England and Wales, religious education has long been seen to have a role in citizenship education (Jackson 2003, Jackson 2004b, Freathy 2008). Post 9/11, it is seen as playing an important role in aiding pupils to develop the balance of intellectual and social values necessary for living in a pluralistic society, such as tolerance and respect. This suggests that if pupils are able to talk openly across and about differences of opinion in their religious or philosophical points of view, they will then develop into responsible citizens in post-modern inter-cultural democracies. Blaylock writes that 'religious education has, for many years,

aimed to prepare young people for life in a diverse community by promoting attitudes of tolerance and respect towards those who believe and live differently from themselves' (Blaylock 2003, 211).

As noted above, this rationale is becoming increasingly important internationally (section 3 introduction, see OSCE/ODIHR 2007). Recent international studies on the relationship between peace education and religious education have explored it (e.g. Jackson and Fujiwara 2007). Many countries have modified their approach to religious education because of its wider role in citizenship (Jackson *et al.* 2007), but for example: the shift in French education from an avowedly non-religious (*laïque*) stance to the handling of religious material (Blandin *et al.* 2004, Estivalèzes 2005, Estivalèzes 2007, Willaime 2007); the shift from a confessional approach to an intercultural one in Quebec (Ouellet 2004); and the re-introduction of religious education into Turkish schools in the 1980s (Duman 2004). In England and Wales, the Framework recognises this, stating that 'it places special emphasis on...the celebration of diversity in society through understanding similarities and differences' (QCA 2004, 8).

The first striking point is that these kinds of outcomes are not explicitly included as part of learning from religion in the Framework. After all, one might have expected these to be the sort of things that pupils should learn from religion. But, as was noted earlier, the Framework's definition of learning from religion is skills-based, and infused with a summative assessment paradigm. Grimmitt's original term for an aspect of learning has become a label for a cluster of measurable skills, and clearly it is difficult to measure these attitudes and values. These

cannot be attainment targets as currently understood. The Framework however still mentions these qualities. It sets them out as 'attitudes' which 'are essential for good learning in religious education and should be developed at each stage or phase of religious education' (QCA 2004, 13). The list goes on to include 'self-awareness', 'respect for all' and 'open-mindedness' (QCA 2004, 13).

Among academic writers this goal is debated. For some, it is a danger because it distorts the educational value of the subject; Wright considers that it is likely to lead to relativism and theological pluralism:

> On the level of immanence we find a pragmatic approach to truth, in which religion is not taught as an end in itself, but as a tool for encouraging tolerance and mutual understanding in a culturally divided society. The hidden 'truth' here is that religion is no more than a relativistic expression of culture whose primary function is to point beyond itself to our common humanity. On the level of transcendence concerns for social cohesion have often led to the conclusion that ...the only valid theological option is that of a universal theology in which all traditions are equally true (Wright 2003, 287).

However, Jackson is critical of this view, arguing that the aim of promoting social cohesion cannot of itself be shown to lead to relativism or pluralism. He also rejects the idea that religious education could simply be about the achievement of these aims. It is through the development of the critical skills identified above that pupils come to develop the appropriate attitudes, but, 'it would equally be a

mistake, however, to assume that understanding and knowledge necessarily foster tolerance: racists can be well informed' (Jackson 2005, 11). Thus, just as use of religious education to promote social cohesion will not inevitably lead to relativism or theological pluralism, so it will not *inevitably* lead to the inculcation of these civic values.

It would be wrong however to assume that Wright rejects the need for a sense of others. He identifies the virtue of 'receptivity', which 'requires students to become sensitive not merely to their own thoughts, feelings, narratives and meanings but also to those of the other individuals, groups and narratives they encounter' (Wright 2004, 224). Thus, even he wishes pupils to be able to deal with some of the challenges of living in a pluralistic society.

This difference in treatment between the assessment levels and the attitudes and values in the Framework, and the preceding policy documents can be seen as a wider tension between two classical functions of schooling. Parsons identifies the two main functions of schooling as socialisation into shared cultural norms and values and differentiation of the labour force by merit (Parsons 1959). Summative assessment plays a vital role in the second function. The problem is whether it reduces the role of the first function.

As far as the context of practice is concerned, it is difficult to comment on how teachers interpret the Framework's guidance on this as it is so recent. However both Hayward and Everington suggest that teachers are likely to interpret it widely. Responses in Hayward's research suggest that teachers interpret the

attainment target in a range of ways that include both reflective skills, the development of attitudes and values, and an appreciation of significant religious figures, for example: the development of '[pupils'] own beliefs and opinions'; 'how to treat people sympathetically/empathetically'; 'the importance of the person of Jesus, respect for his teaching' (Hayward 2007, 81-82). Similarly Everington's research into trainee teachers showed that they valued being able to interpret this target freely (Everington 2007).

What does this review of the place of attitudes, values and dispositions reveal? Firstly, it shows us that there is a lack of alignment between the formal attainment targets and what for many are important goals in the subject. It is also clear that this is an area where, in the context of influence, voices are divided. Moreover, in the context of practice, there is some evidence to suggest that teachers are both interpreting the texts in a writerly way, and also regard the freedom to do this as important.

3.4 Self-assessment in religious education

Finally, where is self-assessment located in the debates described above? Grimmitt himself does not discuss pupil self-assessment at all. He argues for a range of assessment strategies, including testing where appropriate, and he particularly highlights the use of profiling by teachers 'for the regular assessment of the *individual* pupil's progress' (Grimmitt 1987, 254 – author's emphasis; see also Kincaid 1991,123), which echoes the notion of 'ipsative' assessment (Black 1998).

Self-assessment is mentioned in the Westhill and FARE projects. The Westhill project suggested that:

> It is this self-monitoring element which is the most appropriate medium for carrying out those broad formative assessments relating to pupil development in the areas of sensitivities, patterns of awareness, personal qualities and attitudes, beliefs and values. These are central to RE. (Westhill College 1991, 50)

Similarly, the FARE Report stated that 'pupil self-assessment contributed to pupil development in one of the essential areas of RE, the ability to reflect on experiences and personal response to these, and to make judgements about meaning' (Copley *et al.* 1991, 86).

Further, the point was reiterated by QCA in its Guidance: 'not all assessment in RE consists of tests or external judgement. Self-assessment is particularly useful in some aspects of learning from religion' (QCA 2000, 23). In my earlier article (Fancourt 2005a), I explored some of the issues with this claim. In particular, I pointed out how self assessment of learning about religion was often equated with learning from religion. In practice, there were few examples of self-assessment of learning from religion, i.e. pupils' consideration of whether they were better at reflection or evaluation. In particular, there were no examples of how pupils could self-assess the more affective and personal aspects of the learning from religion (Fancourt 2005a, 122-123), e.g. 'am I more respectful of others?' This is shown diagrammatically in Figure 1.

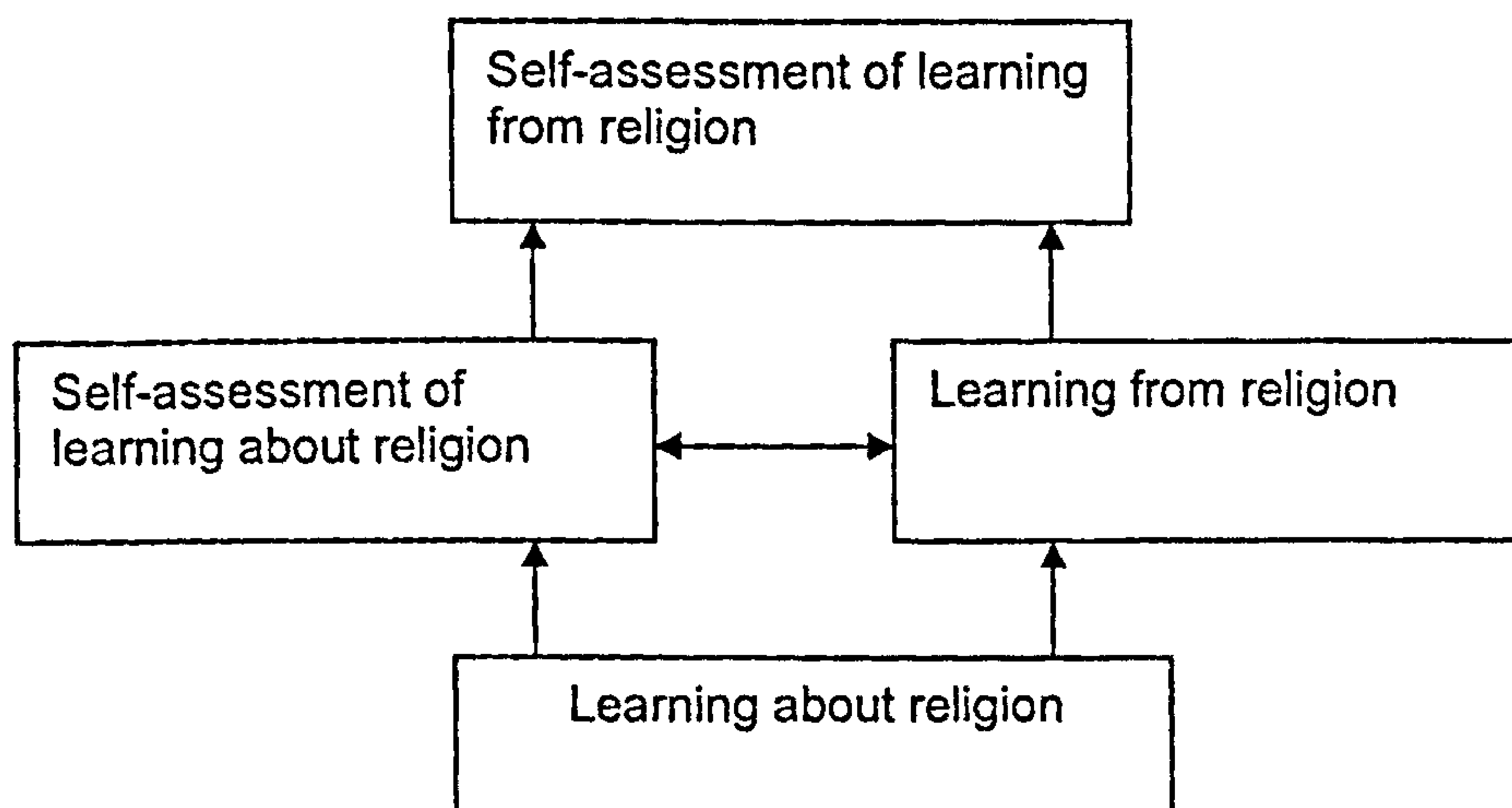


Fig. 1: Levels of reflection in RE

Nevertheless, parallels can be drawn between self-assessment and reflective elements in various pedagogies. As noted above, Grimmitt's description of 'learning from religion' includes a number of questions that, though intended existentially, can be viewed within the context of self assessment: questions of 'self-adjustment', e.g. 'how ought I to change? What do I have to do in order to change?'; and 'questions of self-illumination', such as 'how can I judge my progress?...have I learned anything more about myself?' (Grimmitt 1987, 228). It is however very surprising that he does not develop this idea himself.

There are also strong parallels with Jackson's interpretive pedagogy, despite his lack of explicit attention to assessment. He distinguishes three 'aspects of reflexivity':

- Learners re-assessing their understanding of their own way of life (being 'edified' through reflecting on another's way of life).

- Making a constructive critique of the material studied at a distance.
- Developing a running critique of the interpretive process – being methodologically self-critical (Jackson 2004b, 88, see also Jackson 2000b, 134).

The first two aspects describe how pupils reflect on and evaluate their own views and the views of others. In the first, the use of the term 're-assessing' is arresting, but not ultimately significant, though the use of the term without considering its connotations is surprising. However, the third aspect is particularly striking as a description of the way that pupils can self-assess their learning, both about different religions and about the other reflexive processes of edification and critique. Moreover, as has been highlighted above, Jackson considers that religious education can contribute to the development of vital social values in a pluralistic democracy (Jackson 2004b, 161-180).

3.5 Policy and pedagogy in religious education: a general question

There is clearly a lively discussion about learning goals and attainment targets in religious education. What is the inter-relationship between learning from and learning about religion? What is the relationship between these and social values? How are these learning goals aligned with the current attainment targets? These issues can be summarised in a second general question:

- Are the current attainment targets adequately aligned with all the learning outcomes in religious education?

The first general question was on the classroom factors that affect self-assessment (see section 2.6). For clarification, these two general *questions* will help provide the answers to the general *issue*, described in chapter 1:

- What is the nature of self assessment in religious education?

This issue therefore draws together both the issue of pupils' understandings of their learning with the issue of the goals of religious education. It is important to find out how pupils conceptualise the subject's pedagogic aims, and their progress towards them. This throws into question a wider set of fundamental issues underlying on the one hand religious education curriculum and pedagogy, and on the other the value and purpose of assessment, understood in its widest sense. Religious education is probably unusual in having so central a reflective component, but whether and how they can be harnessed together is a challenge, theoretically, for research, and in practice.

3.6 Conclusion

This chapter contributes to the significance of the research by setting out the academic and practical issues surrounding assessment and attainment in religious education. It has reviewed the development of the current attainment targets in religious education: 'learning about religion' and 'learning from religion'. In particular the changes in three policy documents have been charted: the Model Syllabuses; the Guidance; the Framework. Some of the pedagogical

arguments that surround this have also been set out, and the tensions between pedagogy and policy, notably in the work of Grimmitt, Jackson and Wright. The explicit and implicit potentialities for self-assessment in various pedagogies and policy documents have been described, notably interpretive religious education. Finally, a second general question has been put forward.

The chapter contributes to the research's distinctiveness in that it uses policy scholarship to analyse the place of the two attainment targets, in particular Bowe *et al.*'s theory of three contexts (influence, text production and practice) and their use of Barthes' poststructuralist theory of interpretation. It has shown how Grimmitt's distinction between 'learning about religion' and 'learning from religion' has been adopted and adapted in the contexts of influence, policy text production and practice. It also contributes to the originality of the research because it is now possible to inter-relate the generic research on assessment for learning with pedagogy and policy in religious education, through the two general questions that break down the general aim. At this point, methodology should be addressed, so that the issues of quality and rigour can be considered.

Chapter Four

The virtues of practitioner research: methodological issues

Having outlined the issues arising from the theoretical and research literature on assessment and religious education, the next two chapters explain how this research was planned and carried out. A search has not identified any examples of research specifically into self-assessment in religious education. The research methodology therefore needs careful consideration, particularly as this was research in my own teaching. Is, however, a teacher's practical focus compatible with academic research? As Pring asks,

> Does the concentration upon the uniqueness and distinctiveness of an 'educational practice' preclude the relevance of research as that is normally understood? (2000, 134)

Similarly, Cohen, Manion and Morrison ask whether such research 'is an optimistic way of ensuring that research impacts on practice for improvement, or whether it is a recessive hybrid' (2000, 241), and Hammersley asks if it is 'a contradiction in terms'? (2004, 165). The nature of such research therefore needs to be explored, and the different varieties, such as 'action research', 'practitioner research' and 'teacher research' distinguished.

The methodology proposed for this investigation is a virtue theory research paradigm for practitioner research, using qualitative, ethnographic methods. In this chapter the first two elements will be explained and justified. In the first

section, the argument for virtue theory as a research paradigm is made, rather than the traditional paradigms of positivism, interpretivism and critical realism. In the second section, the nature and quality of practitioner research are considered. The next chapter looks the appropriateness of quantitative or qualitative methods, particularly ethnographic approaches. It then identifies the *research* questions, as research-orientated refinements of the *general* questions, and the last section explains and justifies the research design and ethics.

4.1 Virtue theory as a research paradigm

The importance of intellectual virtues in educational research has been recognised by Pring (e.g. Pring 2000), and Oancea and Furlong, who have suggested that the use of virtue theory would 'make progress towards recapturing a cultural and philosophical dimension of research assessment that had been lost in recent official discourses' (Oancea and Furlong 2007, 119). They use Aristotle's terms to suggest that three overlapping domains can be identified: *episteme*, or demonstrable knowledge, e.g. academic research; *techne*, or technical skill, e.g. using research-based knowledge to solve practical problems (Oancea and Furlong 2007, 126); *phronesis*, or practical wisdom (often translated as prudence), 'from this point of view, practice is more than the application of skills; it involves deliberation about ends and reflective choice, and so is ultimately both educational and social' (Aristotle 1953, Oancea and Furlong 2007, 126). Comte-Sponville calls it 'good sense, but in the service of good will' (Comte-Sponville 2003, 32).

As a teacher researcher, my research is located within the 'phronetic' domain; however it pays attention to the other two domains. How therefore do these domains interrelate? Following Oancea and Furlong, it is argued that this is best conceptualised in terms of different virtues. These different virtues are identified, especially drawing upon current developments in virtue epistemology, which prioritize the epistemological centrality of virtues (e.g. DePaul and Zagzebski 2003). Such an approach can overcome some tensions in research; this goes beyond Oancea and Furlong's argument, to suggest that not only practitioner but all research is best understood in this paradigm.

4.1.1 What are research paradigms?

Social science research is often portrayed as a choice between three research 'paradigms' (e.g. Habermas 1972), which can be briefly sketched, though at the risk of over-simplification. One approach, positivism, treats the social sciences as analogous to the natural sciences; it stems from Comte's 1844 work *Discourse on the Positivist Spirit.* In this essentially empirical method, 'social phenomena are objects and can be treated as objects' (Durkheim 1937, 27[1]). The role of the social sciences is seen as 'discovering the universal laws of society and human conduct within it', and is particularly associated with experimental and quasi-experimental research methods (Cohen, Manion and Morrison 2000, 9). Much of the research on assessment is carried out within this paradigm: for example, Swift and Gooding's analysis of the length of time that teachers wait for pupils to answer questions (1983), or Butler's analysis of pupils'

[1] Translation NF: 'les phénomènes sociaux sont des choses et doivent être traités comme des choses'

responses to different types of feedback (1988). In religious education, an example is Francis and Fearn's analysis of learning styles in 'A' level students (Francis and Fearn 2001).

The second paradigm is a naturalistic or interpretive one. This takes seriously the autonomous nature of people as active, self-aware agents. Its purpose is 'to understand social reality as different people see it and to demonstrate how their views shape the action which they take within reality' (Beck 1979, quoted in Cohen, Manion and Morrison 2000, 20). This paradigm also dates from the mid-nineteenth century, in ethnographic or ethnological approaches to the diversity of human societies being encountered by Europeans around the world (Lienhardt 1964, 3-6). It attempts to analyse the language and meaning of people's world-views and 'values' (Evans-Pritchard 1940, 15), and thereby identify the 'recurrent themes and...main patterns of activity and interconnexion' (Forde 1954, xvii). Within this paradigm there are a variety of related approaches, such as phenomenology or symbolic interactionism (Cohen, Manion and Morrison 2000, 23). Some research on assessment is carried out within this paradigm, e.g. Cowie's analysis of the socio-cultural dynamics of assessment for learning (Cowie 2005), and on religious education in school, e.g. Ipgrave's research on pupils' inter-religious dialogue (Ipgrave 2002).

The third research paradigm is critical theory. In this approach, the research process is explicitly aimed at identifying and reforming political and social inequalities. Habermas, its principal architect, describes it as 'emancipatory' (Habermas 1972), following Marx's suggestion that the task is not merely to

interpret the world but to change it (Marx 1965, 84). In educational research, Griffiths describes it as 'using research for working towards justice, fairness and equity' (Griffiths 1998, 3). In research on assessment, Mejía and Molina's analysis of autonomous thinking styles in classroom conversations is based within this paradigm (Mejía and Molina 2007). In particular, it has had a strong influence on the development of practitioner research and action research, as it envisages the researcher as having a participatory role in the process being studied, for instance Pearce's study of how she dealt with issues of ethnicity as a primary school teacher (Pearce 2004).

Unsurprisingly, the advocates of each paradigm are often critical of others. Positivists may consider that naturalistic approaches miss the reality of the social situation by focusing on the subjective, and are frustrated by the lack of any generalisable principles. They are also sceptical of the effect on data collection and interpretation of the critical theorists' overtly political agenda, which may make it un-objective. Naturalists are wary both of the possibilities of general laws in social phenomena which positivists claim to identify (e.g. Weber 1978) and of ignoring the way that people interpret their social reality; they are also concerned with the effect of a political agenda on the process of the interpretation of people's world-views. Critical theorists argue that the other two approaches ignore 'political and ideological contexts' (Cohen, Manion and Morrison 2000, 28).

What lies behind these tensions? The dispute between positivism and naturalism is often portrayed as a dispute between two different conceptions of the social world. Burrell and Morgan suggest that there is an objectivist

approach, which considers that the generic principles about the social world can be deduced in a realistic, deterministic way. A subjectivist approach is cautious of drawing general principles, considering that people have considerable agency in creating social structures; it focuses on the particular in an idiographic way, taking a constructivist approach (Burrell and Morgan 1979). Part of this is also an ethical concern to take seriously the voice of the people studied; thus Ions criticises behavioural science for 'depersonalisation' of human activity (Ions 1977). The critique by critical theory of other approaches can also be seen as an essentially ethical critique of positivism and naturalism for not attempting to combat injustice. Critical theory is less concerned with the details of research method than with the uses to which it is put, and whose interests it serves.

4.1.2 The strengths of virtue theory as a research paradigm

What is needed for any research is an epistemological approach that provides a conceptual framework for evaluating both the qualitative and the quantitative. This is particularly so, given the demands for a combination of methods: 'the problem becomes one of determining at which points they should adopt the one, and at which the other, approach' (Merton and Kendall 1946, cited in Cohen, Manion and Morrison 2000, 45). However, what is also needed is a paradigm that allows the intertwining of epistemological and ethical concerns, so that the research ethics, in the widest sense, can also be integrated (Winkler 2003).

Virtue theory allows for a range of ethical and epistemological issues to be held together. It draws on virtue ethics, which is well understood as being focused on peoples' qualities rather than the nature and justification of their actions, in a line

from Aristotle through Aquinas to more recent formulations (e.g. Aristotle 1953, Anscombe 1958, MacIntyre 1981).

Virtue epistemology attempts a similar approach, by focusing on truth-bearing qualities of *persons*, rather than on the justifications of their beliefs; thus 'virtue theories in ethics have been described as person-based rather than act-based, and virtue epistemology...as person-based rather than belief-based' (Greco 2004, 1). Intellectual virtues are generically the qualities that underpin knowledge, but can be identified more specifically, for example: memory, eye-sight, open-mindedness, rigour and sincerity. This list includes both attributes (e.g. memory) as well as motivations (e.g. sincerity). A central argument of virtue epistemology, notably Sosa's, is that it resolves the conflict between empiricist and constructivist approaches because it both avoids the difficulty empiricism has in getting beyond an appeal to the senses as the basic justification for knowledge, and it avoids the problem constructivism has of having no clear correlation with what our senses perceive (Sosa 1980, 2003). It would also bypass the vice of scepticism in research (Hookway 2003): how can one be *completely* sure of one's findings?

As a research paradigm, such an approach would allow for a more nuanced understanding of the interrelationship between different kinds of data. It also allows one to incorporate an epistemology of the process of reviewing other research. While this process is seen as essential in all research, its epistemological basis is highly unlikely to be the same as the research methodology. Further, it will often involve reviewing research from a variety of different research paradigms. This is a wider issue of 'trustworthiness' than the

detail of the research methods (Oancea and Furlong 2007, 128), or sincerity (Stern 2006, 104-112, see Fancourt 2008).

Some virtue theorists combine virtue epistemology with virtue ethics. The most comprehensive unified theorist is Linda Zagzebski, who argues that intellectual virtues can be regarded as a subset of moral ones (Zagzebski 1996). She claims that knowledge is in itself *normatively* good, particularly in the sense that it is better than accidentally true belief (Zagzebski 2003). However, one of the strengths of virtue theory is that it recognises that conflict between different virtues can arise. This is widely recognised in virtue ethics, such as between the virtue for teachers of equal treatment of all pupils, and that of compassion in recognising the particular circumstances of a pupil. It can also arise in epistemology; 'thus the intellectually courageous thing to do might conflict with the intellectually humble thing to do' (Greco 2004, 8). Finally, it can arise between ethical and epistemological virtues; the intellectual courage may conflict with the compassion; 'the truth is not always kind' (Pring 2000, 151).

How does virtue theory affect the research structure and process? The difference may not be radical on the surface. For example, in terms of a process of inquiry, Hookway draws up a 'partial' list, as follows

a) Identifying good strategies for carrying out inquires.
b) Recognising when we possess an answer to our question or a solution to our problem.
c) Assessing how good our evidence for some proposition is.

d) Judging when we have taken account of all or most relevant lines of investigation (Hookway 2003, 199)

Conventional issues of methodology remain, for example validity, reliability, research ethics, sampling, bias, and the Hawthorne effect. The difference is that the test for these is not whether the researcher has done the right thing, but whether they can allow *us* (as persons) to be confident both of our own and other people's findings; thus they 'draw our interest primarily for the goods secured through their use' (Sosa 2003, 169).

4.1.3 Balancing two practices

The argument so far may have appeared rather arcane: what difference will it make to the research? The original problem at the start of this chapter was how to balance the research interests with the practice of teaching. So far, it has been argued that research can be seen as requiring virtues. However, teaching can also be seen as rooted in virtues. For instance, Carr argues that the key dimensions of teachers' professional values are found, *inter alia*, in 'aretaic' norms, i.e. virtues (Carr 2006), and Campbell argues that 'the nuances of teaching' reveal such virtues as compassion, firmness, self-control, integrity and commitment (Campbell 2003, 25).

Both teaching and research are examples of what MacIntyre called a 'practice':

> A practice involves standards of excellence and obedience to rules as well as the achievement of goods. To enter into a practice is to accept

> the authority of those standards and the inadequacy of my own performance as judged by them. It is to subject my own attitudes, choices, preferences and tastes to the standards which currently and partially define the practice...the standards are not themselves immune from criticism, but nonetheless we cannot be initiated into a practice without accepting the authority of the best standards realized so far (MacIntyre 1981, 190).

Each practice has its own set of priorities in its standards: 'goods' here means excellences. The priorities of research are different from those of teaching. Researchers do not have to be compassionate or firm in the way that teachers are, and teachers do not generally have to show a 'disposition to search for the truth even when that might be painful' (Pring 2000, 151).

In this research, these two practices are considered as being of equal importance; Afdal describes them as 'as two separate, but potentially related enterprises' (Afdal forthcoming). Academic research is not of more weight than the practice of teaching or *vice versa*. For instance, Bakker and Heimbrock suggest that my paper on the dialogical teaching '*lifts* the reflection of "reflective practitioner" *to the level of* doing adequate research' (emphasis added); while I am grateful for the compliment, this does suggest that research is of more importance than practice (Bakker and Heimbrock 2007, 11; Fancourt 2007). Instead, a better approach is, as Afdal suggests, one that 'opens up possibilities for change that is not "top-down" or "bottom-up", but a result of confrontation and cooperation between the fields of research and practice' (Afdal forthcoming).

Afdal's theory is not a model of practitioner research, but of the interrelationship between the fields. The question is whether these different sets of priorities can be successfully co-ordinated by one individual to allow 'confrontation and cooperation' to emerge. There will be tensions from doing both, such as between adopting what one considers to be best practice with all one's classes and treating a class as a 'control group' in research. However, there are tensions within each practice in any event. The issue is rather how the tensions are negotiated in order both to teach and to conduct research satisfactorily. Afdal, like Oancea and Furlong, labels this 'phronesis' or practical wisdom (Afdal 2007b, Oancea and Furlong 2007, 126). This is the virtue that has to balance these conflicting demands in teacher research, in order both to carry out the research inquiry, epistemically, and to carry out one's role as a teacher, pedagogically. In the next section, the details of practitioner research will be considered in the light of virtue theory, as explained above.

4.2 Practitioner research

In recent decades, models of practitioner research have been developed. The first central point is that such research is,

> carried out by people directly concerned with the social situation that is being researched. In the case of the social situation of the classroom this means in the first place teachers who take professional responsibility for what goes on there (Altrichter *et al.*, 1993, 6).

This involvement raises two issues. The first is whether or not the research should be socially committed: that is to say, whether or not involvement means an ethical commitment to developing social justice for those involved. The second is whether this involvement affects the validity and reliability of the research.

4.2.1 Emancipation or professional development?

In terms of social commitment, some argue for 'emancipatory' action research, which,

> ...seeks to develop in participants their understandings of illegitimate structural and interpersonal constraints that are preventing the exercise of their autonomy and freedom (Cohen, Manion, and Morrison, 2000, 231).

It is no surprise these researchers explicitly base their methodology in Habermas's critical theory, which was discussed above (section 4.1.1; see Carr and Kemmis 1983, Kemmis 1993, Carr and Kemmis 2005). They argue for a 'critical educational science' that aims at 'emancipation from irrationality, injustice, alienation and suffering' (Kemmis 2006, 463). Its roots also lie in Lewin's research to improve the social conditions and expectations of disadvantaged groups (Lewin 1947).

There are a number of criticisms of this approach, apart from the general criticisms of Habermas's work outlined above. It has been criticised for being

too prescriptive in limiting action research to one model; there may be other ways of doing such research (Elliott 1991, 13). It also can be seen to impose a narrow view of 'emancipation' in political terms, and favours an image of self-critical *communities* rather than the individual teacher-as-researcher (Cohen, Manion, and Morrison 2000, 233). Further, the emphasis on effecting change can de-stabilise the quality of the research (Hammersley 2004).

Others maintain that teacher research is rooted in teachers' professional development. Elliott linked it to 'professional learning' (Elliott 1991, 12), as part of curriculum development, not emancipation (Elliott 2005). He drew on the work of Stenhouse, who argued for:

> ...autonomous professional self-development through systematic self-study, through the study of the work of other teachers and through the testing of ideas by classroom research (Stenhouse 1972, 144).

Elliott however did not reject the notion that action research would improve the situation. Rather, he maintained that it was 'with a view to improving the quality of action' (Elliott 1991, 69), and should 'make a significant contribution to the development of a tradition of understandings—a theory—about how to bring about worthwhile educational change' (Elliott 2007, 241). Thus, such research has a 'developmental perspective', and starts from a desire to 'improve a practical situation' (Altrichter *et al.* 1993, 35), without necessarily seeking to resolve issues of social inequality.

This insight has been applied to theories of professional knowledge and reflective practice, particularly the work of Schön, who emphasized the need for professionals to go beyond merely 'reflection in action' to 'reflection on action' (Schön 1983) and Eraut's theory of 'deliberation' (Eraut 1994).

> Action research is based on the theory of reflective rationality and sees the construction of (research) knowledge as integral with the development of action. In this sense, we can understand the process of action research in terms of Schön's analysis of the relationship between professional knowledge and professional action. (Altrichter *et al.* 1993, 207)

As mentioned above, my research has been described as exemplifying reflective practice (Bakker and Heimbrock 2007, 11).

This approach is criticised by critical action researchers. Kemmis uses Illich's distinction (Illich 1973) to contrast 'education', which is emancipatory, and 'schooling', which is socially constraining, and thereby laments the fact that 'action research was harnessed in the service of school improvement of a technical kind – improving the smooth operation of schooling rather than improving the kind of education schools offer' (Kemmis 2006, 469). In particular he considers inadequate:

> ...research that aims only at improving techniques of teaching – classroom questioning or assessment, for example – without seeing

> these as connected to broader questions about the education of students for a better society (Kemmis 2006, 460).

This challenges anyone conducting research on assessment, especially indeed when classroom questioning is part of assessment (see section 2.3.1).

It may be that my research manages to connect the improvement in assessment techniques with broader questions about the education of students, since these are linked to questions about goals of religious education in the development of values. However, this should not be seen as an *essential* part of the research. As Elliott argues in a critique of Carr's and Kemmis's work, 'enlightenment comes from self-critiques that are generated from within the organisation of action as a form of practical discourse' (Elliott 2005, 372); James and Pedder argue that professional development is a vital condition for assessment for learning (James and Pedder 2006). The development of such practical knowledge is of value; even if one were committed to social justice, there would still be a need to improve the *techniques* of teaching. Practical questions would still need answering, such as how to enable pupils to read (Wolyie Hussein 2006), to understand the environment (Gough and Sharpley 2005), or to understand set theory in mathematics (Halai 2005). Indeed, two of these examples are from less economically developed countries, Wolyie Hussein's from Ethiopia, and Halai's from Pakistan, where social justice is achieved through sharper practical knowledge.

In terms of virtue theory, this can be seen as a dispute about the priority of values. Does the need to work for social justice override all other

considerations? Kemmis polarises schooling and education, but how would one know whether or not a particular school policy was emancipatory or not? This is a particularly vexed question when considering assessment. As described above, testing can be seen as educationally liberating by some and constraining by others (see section 2.2). One would question the professionalism of a teacher who did not want pupils to improve, but the nature of that improvement, and how one assessed it, might take a number of forms.

Further, it is possible to be committed to social justice but still conduct highly empirical research, which has a value in itself as a process of clarification of the dilemma, or as an attempt to persuade policymakers. To work towards the delineation of a problem is part of the answer. The most striking example of this is Black's and Wiliam's work (notably Black and Wilam 1998a, Black *et al.* 2003) within the ambit of the Assessment Reform Group (Gardner 2006). Daugherty describes how the 1998 paper was a part of strategy by members of the Assessment Reform Group to bring about changes to government policy and thereby affect practice in the classroom. A major theoretical review of quantitative research, with no hint of emancipatory theory, was itself a part of a plan to bring formative assessment into British educational policy because it was both fairer and more effective (Assessment Reform Group 2002, Daugherty 2007). Ethical considerations thus underpinned theoretical academic research.

My research was conducted as a form of professional reflexivity, as 'autonomous self-critical practical discourse' (Elliott 2005, 372), rather than as emancipatory action research. This is to attempt to balance different virtues,

both intellectual and ethical, coherently but creatively. Further, while the issues that Carr and Kemmis raise are pressing, they are not *internally* essential to the research, even if they may underpin both my professional values and my research values. In particular though, it would be important to be alive to the critical issues that assessment presents in the classroom.

4.2.2 Questions of quality

The second issue is epistemological, raised by a teacher conducting research into their own practice. Can they ensure that they are objective, and do their findings have more than personal relevance? This raises the classic research concerns of quality, notably validity and reliability (Cohen, Manion and Morrison 2000, 105).

These two concepts are often considered as features of the research procedures, methods and techniques. For instance, Kumar states that:

> 'Validity' can be applied to any aspect of the research process. It ensures that in a research study correct procedures have been applied to find answers to a question (Kumar 2005, 6).

Further, he defines reliability as 'the quality of a measurement procedure that provides repeatability and accuracy', and he distinguishes both from the researcher's intention to be 'unbiased and objective' (Kumar 2005, 6). For the reasons outlined above, virtue theory would suggest that it is better to view validity and reliability as processes for ensuring personal objectivity and

truthfulness, rather than as a separate category of criteria. They are features of the process by which researchers ensure that they are as truthful as possible, without being overly sceptical (Hookway 2003).

Teacher research has two clear potential strengths. First, it is much more finely attuned to the concrete problems of practice (Oancea and Furlong 2007, 132). Tierney, in her review of research on how teachers' assessment practice changes, sees it as bridging the gap between knowledge-generating educational research and the mediating process of professional development (Tierney 2006, 241). It can be compared to case studies (Adelman, Jenkins and Kemmis 1980, 93)

From an ethnographic point of view, teacher research can be seen as 'insider' research by one of the social participants. Recent studies have questioned the naïve dichotomy between insider and outsider in ethnographic research (Arweck and Stringer 2002). This is firstly because insider research has complemented outsider research; for instance, in classic ethnographic research, Deng's studies of his own people, the Dinka of Sudan, inform and critique Lienhardt's famous studies of them (e.g. Deng 1980, Lienhardt 1961). Secondly, it is because the dichotomy is in itself too extreme, many researchers are both insiders and outsiders, through gender, age or marital status (Reinharz 1997).

Teacher research could also offer greater *ecological* validity, i.e. 'that the situations in the research occur naturally' (Cohen, Manion and Morrison 2000, 110). There is no 'fly on the wall', observing and interviewing; the presence of

the researcher could influence the ways that the pupils and the teacher behave. Their presence may mean that pupils stay silent about sensitive issues, or behave more flamboyantly. It would not involve an artificial programme of teaching for the research, such as Butler's quasi-experimental research on feedback (Butler 1987, 1988). Instead it is the teacher and work that the pupils would ordinarily experience in their usual, 'natural' setting.

If these two strengths arise from the proximity of the researcher to the classroom, the dangers also lie there. To avoid naïvety, the teacher researcher would need to distance his or herself epistemologically from the classroom situation. This can be understood within the wider context of reflexivity:

> Reflexivity requires researchers to monitor closely and continuously their own interactions with the participants, their own reaction, roles, biases, and any other matters that might bias the research (Cohen, Manion and Morrison 2000, 141).

This will include generic research issues such as social class, gender, background, and intellectual bias (e.g. Bourdieu and Wacquant 1992, 39). However the teacher researcher faces the challenge to objectivity by being the key catalyst in the classroom. In terms of research methods, it could be difficult to 'observe' oneself in participant observation, and further it may mean that data from pupils was both invalid and unreliable because they would not respond truthfully.

To surmount this problem, Altrichter *et al.* propose four quality criteria: considering alternative perspectives; testing through practical action; ethical justification; practicality (Altrichter *et al.* 1993, 74-81). First, considering alternative perspectives involves drawing on other people's perspectives, perspectives from other research methods, and from comparable situations. This includes the process of triangulation:

> Triangulation...is about complementarity, and nothing to do with mutual validation. The two observations or methods must be directed at different aspects of the wider phenomenon to be investigated (Gorard and Taylor, 2004, 8).

One can reduce the risk of misinterpretation and bias by applying complementary approaches. While this will not resolve the lack of reliability, since what one finds in one classroom may not be generalisable, it will be comparable with other situations (Altrichter *et al.* 1993, 75). This is in any event true of much ethnographic research in education, notably Willis's famous study (1977) of just twelve working class 'lads'.

The second criterion is testing through practical action. The hallmark of teacher research is in 'developing action strategies and testing them by putting them into action' (Altrichter *et al.* 1993, 77). One common feature of action research, whether practical or emancipatory, is iterativity: the pattern of trialling a technique, then reflecting on it and then trialling it again, and then reflecting, and so on (Elliott 1991, 69-90; Altrichter *et al.* 1993, 6). This, it is argued, is a way of 'scrutinising' the theories reached, as an alternative to conventional

approaches to validity and reliability (Altrichter *et al.* 1993, 74-77). However, while this may be a common feature, it is not an essential one; one can simply investigate a problem or develop a better understanding of an aspect of teaching. Such approaches tend to avoid the label 'action' research, which is now synonymous with the cyclical process, and can be termed more generally 'practitioner' research (e.g. Campbell *et al.* 2004, 83-84, Burton and Bartlett 2005, 38-39, Fancourt 2005b).

The third criterion is ethical justification. As outlined above, the challenge for teacher research is to balance the different requirements of both research and teaching. For instance, it would be inappropriate to carry out research by attempting to teach badly in order to see what the consequences were, and 'covert observation would be incompatible with supporting openness and trust between teachers and pupils' (Altrichter *et al.* 1993, 77). This would also mean that standard research ethics such as confidentiality and consent would need to be included (e.g. British Education Research Council 2003).

The last criterion is practicality. The research should be both compatible with 'the teacher's main responsibility for teaching', but should also enable other teachers to build upon it on their teaching as well (Altrichter *et al.* 1993, 80). Similarly, Oancea and Furlong suggest that technical research should have a 'concern for enabling impact' and 'operationalisability', and research in the phronetic domain should be both plausible and salient to teachers (Oancea and Furlong 2007, 130-132).

4.3 Conclusion

This chapter contributes to the significance of the research by addressing an important issue for both research and practice: the nature and status of practitioner research. Is it genuine research? Is it possible to combine it effectively with practice? The context of these disputes is set out, in outlining the three traditional research paradigms: positivism, interpretivism and critical theory. It is then argued that this dispute can be analysed within the wider epistemological and ethical framework of virtue theory (e.g. Sosa 1980, DePaul and Zagzebski 2003). This framework is then used to analyse the argument between emancipatory practitioner research (e.g. Carr and Kemmis 1983) and research as professional development (e.g. Elliott 1991). Issues of quality are then considered, notably the place of iterativity in practitioner or action research, and ethical issues.

Moreover, the chapter is different because it resolves these issues through the sustained use of virtue theory. Other philosophers of educational research have argued for the use of virtue theory, notably Pring (2000), and Oancea and Furlong (2007). However, they do not use it in a radical way to underpin all research epistemology and ethics, and they do not draw on more recent developments in virtue theory.

This approach sets the framework for judging the issues of quality in the research. In particular, the strengths and weaknesses of practitioner research are reviewed, notably the issues of validity, reliability and ethics in the light of

the dual roles of teacher and researcher; this raises issues of reflexivity. The more detailed research design of this framework is explored in the next chapter.

Chapter Five

From research theory to research practice: design and method

How should this research be designed? The challenges for my research design are firstly that there are no other examples of classroom research into self-assessment, so there is no 'blueprint' (Cohen, Manion and Morrison 2000, 73). Secondly, as practitioner research, it should balance the demands of both research and teaching, and also maintain the appropriate standards of quality, as discussed in chapter four.

A pilot study was an important influence on the final main research design. Strictly chronologically therefore, the pilot study should be described first, and then the methodology for the main research, but this would alternate between methodological and substantive issues. It is simpler therefore to consider all the methodological issues together (in chapters four, five and six), and then analyse the substantive issues from the pilot study in chapter seven and from the main research in the following three chapters.

In this chapter, the first section considers the choice of methods for practitioner research, qualitative or quantitative, arguing for qualitative ethnographic techniques. After this, the research questions are formulated, as more methodologically specific versions of the general questions. Next, the research design is explained, and finally ethical issues are re-considered.

5.1 Qualitative or quantitative?

Should research on assessment be quantitative or qualitative? Almost all of the research evaluated by Black and Wiliam was quasi-experimental, showing clear measurable differences in attainment by pupils who experienced assessment for learning (e.g. Butler 1988, Fontana and Fernandez 1994, in Black and Wiliam 1998a). A common approach in teacher research is to trial new techniques and what difference they make to pupils' attainment, (e.g. Harris and Foreman-Peck 2001, Halai 2005).

This approach would have three problems. Firstly, as has been explained in the chapter three, the current two attainment targets are not aligned with the range of possible learning outcomes in religious education; many consider that they do not adequately describe the intellectual learning processes, and further they do not describe criteria for the development of attitudes and values. In other words, they lack content validity, because the instrument would not 'fairly and comprehensively cover the domain that it purports to cover' (Cohen, Manion and Morrison 2000, 109). While it would be theoretically possible to develop a similar classroom-based set of levels for attitudes and values, there would be a considerable question mark over the validity of such an instrument (see Narvaez *et al.* 2004).

The second problem arises from a potential conflict between the role of teacher and the role of researcher. In the examples given above of quasi-experimental research, the teacher could adopt a formative role while the researcher could administer the research 'test'. The roles are separate. However, there is a tension in explaining to pupils that they are to be involved in research that is designed to

move away from testing, and then to administer a research 'test'. It could reinforce an ego-focused view within the pupils, whereas as a teacher I would be trying to mitigate this. Administering the research instrument could therefore negate the wider intentions of the research. Moreover, as the teacher and researcher, there is a danger that the instrument lacks construct validity because it would be difficult to judge whether my conception of the values and attitudes agreed with other constructions. This was not a problem for Harris and Foreman-Beck because they used the 'A' level examination as the research instrument, so the examiners provided inter-rater reliability (Harris and Foreman-Peck 2001).

Thirdly, the issues of generalizability and reliability need to be balanced with the practicalities of carrying out the research while teaching; too small a sample would make the results statistically unreliable. It would be possible to carry out quantitative tests on a whole class, or on all the classes that I taught, but it would be difficult to sustain a focused development of self-assessment with all my classes in this way.

There are therefore strong reasons for not adopting a quantitative approach. By contrast, a qualitative, ethnographic approach would focus on pupil perceptions of their learning, which is precisely what self-assessment sets out to achieve. O'Grady, when researching the impact of interpretive religious education through action research, adopted ethnographic methods and asked the pupils to be 'co-researchers' (O'Grady 2003, 214). Secondly, it would mean that the roles of teacher and researcher are not in conflict, particularly because discussions are an important feature of religious education lessons, so further research interviews can follow from them.

There are other strengths to qualitative ethnographic studies. Much research, notably from University of Warwick, has shown how detailed studies of the ways that individual children understand their religious belonging can have a significant impact in developing pedagogical approaches (Nesbitt 1991, Nesbitt and Jackson 1995, Nesbitt 1998, Østberg 2000, Jackson 2000b, Østberg 2003). In particular, Julia Ipgrave's research on pupil dialogue in the classroom has contributed greatly to the development of dialogical pedagogy (1999, 2001, 2002, 2003, Ipgrave and McKenna 2007).

There are also drawbacks to this approach that need to be recognised. A major weakness is the lack of any external indicator of progress. The second is the risk of lack of objectivity as a result of being both teacher and researcher, and the possible effect on the pupils of this combination of roles. These would need to be addressed as much as possible when formulating the research questions and in the research design.

5.2 From general questions to research questions

So far, the broad issues presented by the research literature have been explored, both in assessment and in religious education, and the broad contours of the research methodology. The overall general aim (in chapter one) was:

- What is the nature of self-assessment in religious education?

This however was broken down into the two separate general questions from the review of the two fields of literature. The general question on assessment was formulated in section 2.7:

- What classroom factors affect self-assessment?

The review of the literature on religious education led to the second question, in section 3.5:

- Are the current attainment targets adequately aligned with all the learning outcomes in religious education?

These general questions need to be translated into *research* questions, in the light of the decisions about methodology, and particularly a focus on pupils' perspectives.

It should be recognised that the precise formulation of these research questions arose out of a pilot study that was carried out before the main research; it would be misleading to think that the process was as smooth as a sequential reading may suggest. The pilot research aim was different, as it focused on the techniques themselves: which self-assessment techniques are most suitable for the different aspects of learning in religious education? It was also broken down into four research questions that tackled the individual techniques. The reasons for the changes are discussed in section 6.2.4 below.

The first general question, on self-assessment in the context of the classroom, can be related to pupils' perceptions of their learning. Furthermore, it can be divided into three sub-questions, which relate to the specific self-assessment techniques, issues of motivation, and other factors:

- How can we understand self-assessment in the context of pupils' own ideas about their own learning and the classroom setting?

  a) How do pupils' views on different self-assessment techniques vary?

  b) Do pupils become more motivated as their ability to self-assess develops?

  c) What else affects pupils' identities as learners?

At one stage, during the pilot study, a sub-question on gender was considered. However, the small numbers of pupils involved, the introduction of another variable, and the implications for further understanding of the issues in the relevant literature made this impractical, even though it is important in assessment (e.g. Gipps and Murphy 1994, 28-64).

The second general question, about the current attainment targets in religious education, is also not framed to focus on pupils' perspectives. It is therefore helpful to re-frame it, and to break it into two sub-questions that separate out what the pupils consider to be the goals of the subject from whether they think have made progress towards them, because, as any footballer knows, seeing the goal is not the same thing as progressing towards it:

2. How do pupils consider that self-assessment helps their progress in religious education?

a) What do pupils consider to be progress in RE?

b) Do pupils think that they have improved at this?

In this form, the question will allow the pupils' voice to be considered on these issues. The structure of the questions can be set out as a table; see over.

Table 1: Question structure in the research

General aim	What is the nature of self-assessment in religious education?	
	Assessment for learning strand	**Religious education strand**
General Questions	What classroom factors affect self-assessment?	Are the current attainment targets adequately aligned with all the learning outcomes in religious education?
Research Questions	How can we understand self-assessment in the context of pupils' own ideas about their own learning and the classroom setting?	How do pupils consider that self-assessment helps their progress in religious education?
Research sub-questions	a) How do pupils' views on different self-assessment techniques vary? b) Do pupils become more motivated as their ability to self-assess develops? c) What else affects pupils' identities as learners?	a) What do pupils consider to be progress in religious education? b) Do pupils think that they have improved at this?

5.3 Research design for practitioner research

At this point the research design can be drawn up, so as to 'conceptualise an operational plan to undertake the various procedures and tasks required to complete [the] study' in order to ensure objectivity and accuracy (Kumar 2005, 84). There are four dimensions to this in practitioner research, considered in turn below: identifying the precise nature of the implementation, or the 'action steps' (Elliott 1991, 76); data collection, or 'gathering evidence' (Elliott 1991, 77); data analysis (Altrichter *et al.* 1993, 119); developing further action strategies (Altrichter *et al.* 1993, 153).

5.3.1 The action steps

For this research, the action steps could have either been related to self-assessment techniques or to issues directly related to pedagogy. To change both areas would create too many variables in the research, and require an impractical degree of planning in terms of my teaching. I decided to keep the main content and format of the lessons the same, but to introduce various self-assessment techniques. My pedagogical approach inevitably affected how this developed; it was implicitly interpretive and dialogical (e.g. Jackson 1997, Ipgrave 2001, Jackson 2004b, Fancourt 2007), though some activities were drawn from other approaches, e.g. the use of empathy (Fancourt 2003a, 2003b).

What were the appropriate self-assessment techniques to utilize? One issue with which this research is concerned is the extent to which pupils can self-assess the development of their attitudes and values; therefore, a variety of self-

assessment techniques would be needed that would enable them to do this. Further, there is concern about the negative effect of using levels or marks, even if they are used with comments, as described above in section 2.2 (e.g. Black and Wiliam 1998a, 49); therefore more techniques were needed than simply those that explain the attainment levels in the Framework (QCA 2004, 36-37). Thirdly, techniques would be needed that were compatible with the learning activities in the lesson. A combination of the following four activities was practical.

First, 'end of lesson reviews':

> At the beginning of the lesson, one student is appointed as the 'rapporteur'. The teacher then teaches a whole-class lesson...and finishes fifteen minutes before the end...The student rapporteur then gives a summary of the main points...and tries to answer any remaining questions that the students in the class may have. If he or she can't answer the questions, then the rapporteur asks members of the class to help out. (Learning how to Learn 2002)

These would allow pupils to self-assess their learning each lesson, against the learning objectives, which will be clearly stated each lesson. I had always stated the lesson aim; the addition to my practice was the self-assessment element.

Secondly, 'traffic lights':

> The teacher identifies a ...number of objectives for the lesson...At the end of the lesson, the pupils are asked to indicate their understanding of each objective by a green, yellow or red circle on their work, according to whether they feel they have achieved the objective fully, partially or not at all. (Learning how to Learn 2002 (see also AAIA, North East Region 2003, Copley *et al.* 1991))

This description of the technique is similar to the 'end of lesson review'; however, it is possible to use the traffic-lighting process for a review of an entire module rather than just one lesson. Of itself, it does not provide 'feed-forward' (Sutton 1995, 74); it could however be applied throughout a module to encourage this. An example is in appendix 1, which also included two targets that the pupil set for himself, and see also appendix 2, second section.

A third technique that explicitly deals with the current attainment targets and levels would be appropriate. One such technique is the use of understandable criteria to assess their own work and make improvements, and can be combined with peer assessment of other pupils' work (AAIA North East Region 2003, 21-24, DfES 2003). For instance, OFSTED state that in religious education,

> Some of the most effective use of levels takes place where pupils are provided with simplified versions of the level descriptions. They are encouraged to use these to assess their own and other pupils' work, set targets for improvement and discuss the reasons for their judgement. (OFSTED 2003, 6)

However, it should be noted that the view that assessment for learning is essentially concerned with these processes has been criticised for being to narrow (Daugherty and Ecclestone 2006, 154). Appendix 3 is an example of a simple guide to the criteria, referred to as the 'level-path' sheet, on which the pupil has identified their current attainment.

Finally, discussion is a common activity in religious education. This could be extended to include not simply the religious, philosophical or ethical questions that naturally arise, but can also include wider issues of learning and progress. This can be seen as a form of questioning (AAIA North East Region 2003, 14-16), which has been 'more recently relabelled as *dialogue*' (Black and Wiliam 2006, 14, authors' italics). This has echoes of dialogical approaches to religious education (e.g. Ipgrave 2001, Ipgrave 2003, Jackson 2004b, Fancourt 2007). Indeed, the Westhill project group suggested that teachers can listen to discussion as a way of assessing students (Westhill College 1991, 54). In appendix 2, an example is given of a pupil's initial views before such a discussion.

When should these techniques be applied? In order to see how students engage with the processes, it is prudent to let them engage with them over a long period. The research was not on all aspects of pupil learning but on one aspect of assessment alone; however, it is possible that in one lesson, only one self-assessment technique would be used, while in another, all four could be used. In particular, having only one lesson a week with many classes, their development of the techniques could take some weeks, or have no effect (Fontana and

Fernandez 1994, 415). A basic timeframe of a term was therefore suitable and practical.

With what classes should they be applied? There are clear problems with applying them with pupils studying for GCSEs or 'A' levels as the attainment targets are different to the Framework (Keast 2003). However, the school adopts a combined humanities approach in Years 7 and 8 (11 to 13 years), in which one teacher teaches history, geography and religious education to the same class; this approach has been criticised as distorting the subject (e.g. Copley 2005, 115). It would be difficult to disentangle assessment in religious education from geography and history. A Year 9 class (13 to 14 years) would be the most appropriate, since they are taught separately once a week, under a local curriculum that adopted the attainment targets in the Guidance (Oxfordshire County Council 1992). For the pilot study, this would be over one term, and for the main research it would be over a year.

5.3.2 Data collection

The focus of the research was on pupils' perspectives on how self-assessment helps their progress in religious education, and needed to gather the pupils' honest, authentic views, allowing for multiple perspectives, and changes to points of view. The main source of data would therefore be interviews.

There is currently popular support for recognition of 'pupil voice' in schools (e.g. Flutter and Rudduck 2004), though it has long had a place in ethnographic educational research (e.g. Willis 1977). Recent demands are for more

consultation with pupils about teaching and learning *in the classroom*, which has received less attention (Fielding 2007). Self-assessment is about the pupils' ability to reflect on their learning, and the teaching that they receive; if this is to be taken seriously, then pupils' views on whether and how they can develop that ability should also be taken seriously. My research is therefore 'compatible with the educational aims of the situation being researched' (Altrichter *et al.* 1993, 77).

There are however two challenges. First, there is the risk of pupils not speaking freely because my role as teacher inhibits them; Lewis asks that 'teacher researchers consider carefully whether or not the children's inter-relationships or his or her relationship with them distorted responses' (Lewis 1992, 417). Cohen, Manion and Morrison (2000, 287) suggest that a difficulty in interviewing pupils is how to 'get the children's teacher away from the children'. This can be reduced through group interviews, since pupils will be with their peers rather than just with me; this is the approach adopted by Ipgrave (2002) and O'Grady (2003, 2006b). Altrichter *et al.* suggest that 'the group interview is the more normal situation for students' (1993, 103).

There are dangers, such as the effect of group dynamics (Burton and Bartlett 2005, 113), and Altrichter *et al.* warn that there is the ethical danger of pupils ridiculing others (1993, 103). This needs to be monitored, but it is also linked to group size. Four pupils per group, following Ipgrave (2001) and O'Grady (2003) seemed practicable, though Lewis suggested six or seven (1992, 417). In the pilot study, there were four interviews of four students in one term, to test the collection and analysis procedures. However, it would not be practical to interview a whole class in this way; that would require eight interviews each term.

To ensure that it remained practicable, only four groups would be possible, in both the pilot and the main research. This is clearly a compromise between practicality and rigour.

The second danger is reactivity (Kumar 2005, 97): respondents will react differently because of their participation in research. The best known example of this is the Hawthorne effect, when participants behave differently because they are being observed (Kumar 2005, 120). In this case, by being asked questions in interview on self-assessment, pupils may come to reflect differently on the process of self-assessment in the classroom. The research and interview process can be seen to set up six levels of reflection:

(i) Pupils self-assess their learning about religion

(ii) Pupils self-assess their learning from religion

(iii) Pupils self-assess the development of their attitudes and values

(iv) Pupils are interviewed about their self-assessment of learning about religion

(v) Pupils are interviewed about their self-assessment of their learning from religion

(vi) Pupils are interviewed about their self-assessment of the development of their attitudes and values

The danger is that by getting the pupils to engage in the last three types in the interviews, their view of the first three types may alter. The findings may not be a record of what the pupils think of the self-assessment techniques that they have experienced, but rather may be a record of what they think of self-assessment techniques that they have experienced in the light of the interview questions.

This danger is unavoidable, but should be recognised in the analysis stage. It can however be mitigated by ensuring that the balance between lessons and interviews is firmly weighted in favour of the lessons, so that the majority of the pupils' experiences is in them, and not in the interviews. Given that the basic time-frame for the trialling of the techniques was termly, the interviews would also be termly; appendix 4 sets out the plan for the teaching and researching in the main research. The pilot study also provided the opportunity to develop my interviewing skills.

In order to consider alternative perspectives, some other form of data would be useful. There are two which contribute to the interview data. The first is a research diary, particularly to record the details of the lessons in which the techniques were being applied. These draw on a long-standing ethnographic tradition (e.g. Malinowski 1982), and in practitioner research they are vital as a record of what, as a teacher, I was attempting to teach, as well as what, as a researcher, I observed. It is a record of participant observation, as discussed above in section 4.2.2. Clearly however, this record cannot be made during the lesson because I could not both teach and record the behaviour of the class, but would afterwards.

The second approach was to use questionnaires for the whole class. The issues that arise in the interviews could be compared with the whole class, especially as a way of suggesting avenues for exploration in the interviews. This data was not for thorough quantitative analysis, particularly for the reasons described above in relation to the design of research instruments (section 4.5), and because the sample size is small. A simple Likert-scale questionnaire would be used, to provide an overview of pupils' perspectives (Cohen, Manion and Morrison 2000, 253). This is discussed in the next chapter in relation to the pilot study (appendix 5), however it was in fact abandoned in the main research, as discussed in chapter six.

5.3.3 Data analysis

How should the data be analysed? One of the problems of transcription is that the nuances of oral language are lost, particularly in a group interview. In order therefore to capture some of these nuances, the conventions for transcribing classroom interactions that were developed by Torrance and Pryor (1998, 106) in their research on formative assessment were adopted (see appendix 6). This safeguards descriptive validity (Maxwell 1992).

In the analysis of transcripts, there is a strong argument put forward in ethnographic research to use only inductive categories, notably in 'grounded theory' (Glaser and Strauss 1967); this means developing themes that emerge form the transcripts themselves, by being 'theoretically sensitive' (1967, 46). However, Altrichter *et al.* suggest that the most useful approach for practitioner researchers is to use a mix of inductive *and deductive* coding categories, thus

'capitalising on what you already know but remaining open to the surprises that the data can contain' (1993, 124).

The coding takes a particular pattern. All transcript material is given at least one code; indeed some sections can be given several, and there could be up to ninety codes (Rubin and Rubin 1995). The codes could then be organised into clusters to create general themes (Rubin and Rubin 1995), while paying attention to 'negative and discrepant' cases (Cohen, Manion and Morrison 2000, 150). Although there is coding software available (see e.g. Barry 1998, Lewins and Silver 2005), my data was coded manually, partially in order to ensure sensitivity to the data.

Beyond the need for theoretical sensitivity by any researcher, there is a particular issue of reflexivity for the practitioner researcher, in that there is a greater temptation to read into the data what one, as a teacher, considers should be there: the apparently objectively emergent codes could covertly be my teaching aspirations in disguise. While some reject the need of external confirmation (Ely *et al.* 1991, 164), the approach adopted is for an independent researcher to code a sample of the transcripts, attempting to use the same codes. This can be seen as a form of 'inter-rater reliability' (Cohen, Manion and Morrison 2000, 116), or a way of 'considering alternative viewpoints' (Altrichter *et al.* 1993, 74). More specifically, in this research, it was a test of 'category reliability' (Asher 1984, 15) through peer-checking, because the independent researcher's analysis is not part of the overall research (Ely *et al.* 1991, 163-164). The pupils themselves were not involved in the analysis, unlike Leitch *et al.*'s research (Leitch *at al.* 2007, 469-472).

The final check on the validity of the data is to analyse it in two ways. First, it is considered in relation to the research questions and the two fields of assessment, in chapter seven, and religious education, in chapter eight. Then, the data is re-contextualised more biographically in chapter nine, as a series of case-studies of individual pupils, in which their views across both areas are brought together. This draws on Filer and Pollard's analysis of one pupil's experience of assessment throughout primary school (Filer and Pollard 2000, 37-52), but is considerably more focused on the micro-level of their experiences over a year in one subject. This is not a perspective from another research *method*, which Altrichter *et al.* suggest (1993, 76), but is an alternative, complementary, perspective on the data (Gorard and Taylor 2004).

5.3.4 The lack of action cycles

As noted above, it is a common feature of action research to have a series of action cycles (section 4.2.2), because it is regarded as an important criterion of quality; Elliott argues for at least three cycles (Elliott 1991, 85, see also Altrichter *et al.* 1993, 77, and in research on religious education, O'Grady 2006b). However, there were no formal cycles in my research. This is because of a practical problem in collecting and analysing the data. It would be important to conduct the interviews at a stage when the pupils have experienced a range of self-assessment techniques, i.e. towards the end of each term, with the first interviews in December, and the second in March. The four interviews, of forty to forty-five minutes, would then be transcribed, i.e. 2½ to 3 hours of interview before coding and analysis, and then formulating the next action steps; it would be impossible to achieve this before the next term, i.e. by January and April.

While it would be theoretically possible to halt the process until this had been completed and then restart it, the issue would be what I would teach and how I would assess in the meantime, and how this interruption could affect pupils' attitudes.

My research therefore is a longer term but more static study of pupils' attitudes to and perceptions of the process of being asked to 'renegotiate the teaching contract' (Perrenoud 1991, 92). It can be viewed therefore as a detailed first action cycle, from which later research can develop, or which potentially others could apply 'thus broadening the knowledge base of the profession' (Altrichter *et al.* 1993, 207). It is 'practitioner research' (e.g. Campbell *et al.* 2004, Barton and Bartlett 2005; in social work, Fuller and Petch 1995), used as a more general term to describe research by anyone into their own practice, rather than the more specific 'action' research which would also require action cycles.

Nevertheless, some degree of iterativity is useful in ensuring confidence in the research. Therefore the same techniques should be repeated in each of the three terms, and the pupils interviewed at the end of each term, so that the picture of their views is more reliable. This will also enable individual patterns and changes to be more apparent; it will be easier to see if different pupils react in different ways. It is less a description of improvement, and more a portrait of how pupils respond over time to a particular intervention. However, in the later chapters analysing the data, the letter of the group and the number of the interview will be given, as well as the pupil's pseudonym, so that it is clear when the comments were made.

5.4 Ethical justification

Finally, the ethical underpinnings of the research should be reviewed (British Educational Research Council 2003). As explained above, the practice of research is rooted in the practice of teaching, and has to be compatible with this, such as standards of honesty, for example, a lack of deception (Cohen, Manion and Morrison 2000, 63, Pring 2000, 148). This influenced the decision not to carry out any quantitative analysis in the form of testing of pupils' attainment in the research (Altrichter *et al.* 1993, 77), and the decision to adopt ethnographic techniques. However, there are particular considerations in the collecting of data from pupils.

The first is informed consent, where the nature and purpose of the research is made clear (e.g. Burton and Bartlett 2005, 29, Kumar 2005, 212): the school's consent, the pupils' consent, and their parents' consent. Further, the pupils would always have the right to withdraw from the research – though of course they could not withdraw from the lessons. They would receive copies of the transcripts 'to think over what they have said' and to respond or comment (Altrichter *et al.* 1993, 78). They would also receive a summary of the initial findings.

This is linked to the issue of fair representation, particularly in ethnographic research (see Said 1978). The use of detailed transcription conventions safeguards this and the opportunity for pupils to challenge the transcripts in terms of 'relevance and accuracy' (Cohen, Manion and Morrison 2000, 68). It is also sustained by the secondary stage of analysis, which focuses on particular

pupils, so that their comments and ideas can be understood in the context of their learning.

The last concern is anonymity (e.g. Cohen, Manion and Morrison 2000, 61). The school has agreed to being identified, which is helpful because its involvement in assessment for learning is significant in setting the context for the research. Moreover it is clearly identified in some publications (e.g. OECD 2005). The pupils however are protected by being given pseudonyms.

5.5 Conclusion

This chapter contributes to the research by transparently addressing issues of quality and rigour. It justifies the choice of more qualitative, ethnographic techniques, instead of quantitative analysis, and then explains how more precise research questions were formulated from the general questions.

Next, the main research design has been set out. A critical element is the choice of self-assessment techniques. Four are identified: lesson reviews, module traffic-lights, the formative use of levels, and discussion. Data collection uses largely ethnographic techniques, notably semi-structured group interviews (Lewis 1992), which the pilot study refined. The analysis is through coding, using inductive and deductive categories (Altrichter *et al.* 1993), triangulated through questionnaires, my research diary, inter-rating, and interpreting the data both thematically and biographically. The need for action cycles is considered, before concluding that while these may be one way of ensuring rigour in the research, they are not essential (following, e.g. Campbell *et al.* 2004); thus this is

characterised as *practitioner* research rather than *action* research. Finally, the potential ethical issues that arise from data collection have been identified, such as the reduction of conflict between the two roles of teacher and researcher, and the need for consent.

Throughout this chapter, refinements to the main research as a result of the pilot study have been identified; it impacted on the main research's structure. This added to the rigour of the main research, but it would be wrong to assume that the issues in this chapter were decided before the pilot study took place. The process was more dialectical, as a result of the experience of both conducting the pilot study and methodological reflection. An important part of both the pilot study and the main research was the practicalities of conducting the teaching and researching. In the next chapter, these dual processes are examined and explored in more detail.

Chapter Six

Teaching *and* researching: a review of the process

This chapter describes how the teaching and the researching actually unfolded. It examines whether and how the research was able to match the design, including some of the issues of quality in data collection. First, the school is described, then methodological issues from the pilot study are discussed, and thirdly the practical methodological issues from the main research. Three methodological issues are then considered in detail: the extent to which how being a teacher researcher affected pupils' responses; the extent to which the group interview process affected pupils' responses; the extent to which coding and theory generation can be trusted. Having established these issues, the substantive questions are then addressed in the following chapters.

6.1 Lord Williams's School

The school is a large comprehensive (over 2000 pupils) in the Southeast Oxfordshire market town of Thame, about fifty miles north-west of London. It is split over two sites, about 2 miles apart: Key Stage 3 (11 to 14 years old) on one site, and Key Stages 4 and 5 (14 to 18 years old) on the other. Its catchment includes Thame and a number of surrounding villages. Although nestling between three multi-cultural towns and cities (Oxford, Aylesbury and High Wycombe), less than one percent of its pupils is non-white. Many pupils' parents either work in these towns and cities, or commute to London.

Thame itself has a number of churches; evangelical Anglican, United Reformed church, Roman Catholic and independent Evangelical. There are no other

places of worship. The surrounding villages mostly have an Anglican church; some have another protestant church, e.g. Baptist or Methodist.

The school is a comprehensive, with specialist sports status; there is no selection process. In a school of this size, pupils come with a range of physical and educational needs. Significantly, it has an autistic unit on site, with about thirty autistic spectrum disorder pupils; these pupils are more or less integrated into the main school depending on their needs (see Fancourt 2003b). It is a state school with no religious character, and although there is a legal requirement for schools to hold acts of collective worship, I have never experienced one in over a decade.

The school has been involved in a number of research projects, as described in chapter one, notably KMOFAP, and also some subsequent related projects following on from this. A number of colleagues describe their involvement in this (Black *et al.* 2003, 88-91). It is cited as a case study for implementing assessment for learning across a school by Looney and Wiliam (2005). This background affects the generalisability of this research. These pupils were more than likely to have experienced formative assessment in other subjects, and they were not being asked to change their learning habits completely.

6.2 The pilot study

In this section, the implementation of the action steps in the pilot study is described and then the collection of data is considered. Thirdly, the implications of this for the main research are outlined.

The pilot study had a slightly different research aim and questions to the main research, and the distinction between general questions and research questions

had not been fully articulated. The pilot research aim was:

Which self-assessment techniques are most suitable for the different aspects of learning in religious education?

The pilot research questions were:

- Does traffic-lighting provide help with any aspects of learning?
- Does the use of a review in each lesson help with any aspects of learning?
- Does the formative use of summative assessment criteria help students develop their evaluative skills?
- Does the development of class assessment criteria help the development of empathy?

The aim is typical of action research, directed at the practical issue of 'suitability' of the techniques for religious education, and the questions focus on the particular techniques, 'designed to improve practice' (Cohen, Manion and Morrison 2000, 79).

6.2.1 The pilot study: the class

A Year 9 class (aged 13 -14 years old) was selected. Its key features were:

- It consisted of fifteen boys and fourteen girls. All were white.

- The class included a wide range of attainment. Using key Stage

2 SATS results in the core subjects as a general guide to attainment, three pupils (two girls and one boy) had 33 points; eight pupils had 23 or fewer (two boys and six girls). The national expected median is 27 points.

- Two pupils were identified as having specific learning difficulties (one girl and one boy), and a further six were being monitored (four girls and two boys).

- Three pupils were identified as being 'gifted and talented' from their performance in SATS at the end of Primary school.

- One pupil was a practising Jew. The others, all of British heritage, ranged from atheist to Christian, with varying degrees of commitment.

In terms of the mix of gender and ability, the class was probably reasonably representative of the school, and indeed the national picture. However, the almost complete lack of ethnic or religious diversity should be noted; the pupils did not actively deal with inter-religious or inter-cultural issues on a daily basis.

6.2.2 The pilot study: implementing the action steps

There were the four action steps, in the pilot (see section 5.4.1): the module review sheet (e.g. appendix 1), the development and use of class criteria, the formative use of summative assessment levels (e.g. appendix 3), and the use of class discussion (e.g. appendix 2).

They were mostly straightforward to implement, however, the development of class criteria, in this case for an empathetic piece of writing on Jewish escapees on the *Kindertransport*, did not seem to provide for self-assessment. It did of course aid in the sharing of criteria, but needed to be changed or developed in some way. Further, I noted that there was no regular self-assessment by the pupils; the four techniques that I had chosen did not occur every lesson, which would further reduce the impact of the techniques (see Fontana and Fernandez 1994). In the main research, I therefore decided to replace the development of class criteria with the use of the rapporteur in the main research.

6.2.3 The pilot study: data collection and analysis

The process of data collection, in both the questionnaires and the interviews was also reasonably straightforward. The questionnaires were useful in providing a broad structure to the interviews; it was possible to ask pupils about the general findings in the questionnaires.

The pupils were interviewed in four groups. This was during a science lesson on Friday afternoon over four weeks; each interview lasted about forty minutes. The choice of pupils was negotiated with the science teacher. Three of the groups were mixed, and one was all girls. The interviewees included pupils with a range of educational attainment.

The most striking aspect of the interviews was that pupils were clearly able to talk about their experiences of learning and assessment. Before starting, I was concerned that the interviews would be full of awkward silences, or pupils would simply not understand the questions. This was far from the case.

Transcription following Torrance and Pryor's (1998) conventions proved helpful (see appendix 6). The development of codes was originally deductive, but it became apparent that inductive codes were needed in order to cover all the material, because my preconceptions of what the interviews would include were too narrow, and a range of other issues and concerns emerged (appendix 7, see section 5.4.3). Even despite this however, the use of sixteen codes was too narrow; more refined codes (sixty-nine) were developed and then clustered in the main research (appendix 8).

6.2.4 The pilot study: design modification

There were three significant changes to the main research after the pilot study. First, as noted above, one of the action steps was altered to ensure regular self-assessment every lesson. Secondly, the research questions were re-framed, because there were three difficulties. First, 'suitable' was hard to define. Is suitable the same as effective? Suitable to whom? My research could only answer this question by using the data from the pupils, and while their voice was important in answering the question, other kinds of data would also be relevant.

Secondly, there is a discrepancy between, on one hand, the pitch of the research problem and research questions and, on the other, the themes that emerged for the coding. The themes relate to the wider processes than the research questions allow. Generic issues underpin the pupils' responses, so that a question about whether a technique works depends on their attitude to assessment, to levels, to the social setting, and to understanding the criteria.

Finally, very little data related to attainment in religious education. I originally assumed that the research was simply about the subject. However, the research

questions, and consequently the questionnaires and interviews, did not raise this explicitly. The pupils generally therefore spoke generically about the techniques, such as using traffic-lights; the exception was the material on discussion. The main research problem and questions therefore needed to focus on the understanding, skills and attributes required in religious education.

6.3 The main research

As a result of the pilot research, it was possible to refine and refocus the research aim and research questions, and develop the research methodology, as described above. Did these revisions work? There are three questions:

- What class was selected for the research?
- Was it possible to carry out the action steps as planned?
- Was it possible to carry out the data collection as planned?

6.3.1 The main research: the class

The main research was, like the pilot study, also with a mixed ability group of Year 9 pupils. This created a quasi-iterative element in the research process, more like *action* research (Elliott 1991). But it was different because firstly it explicitly focused on religious education, and secondly the main research was a fuller picture of pupils' perceptions over a year.

The particular class was chosen for practical reasons, because of the timing of the lesson in the school day. It consisted of 14 boys and 16 girls. During the year, one boy left and another joined; neither was interviewed. Seventeen were

interviewed, i.e. just over half. The results of their National Curriculum tests at the end of primary school (in Maths, English and Science) showed wide range of attainment. The median level nationally is 4; one boy had achieved levels 3, 2, and 3 respectively, with a notional reading age when he had joined the school aged eleven of seven years and nine months. Five girls and two boys had achieved three level 5s, and one of these boys had a notional reading age when he joined the school of fifteen years and four months; of these, one boy and two girls were interviewed. A further two boys and three girls had achieved a level 5 in English, which is often regarded as a useful indicator of attainment in religious education (e.g. Wintersgill 2000); three were involved in the research throughout, and one replaced another pupil for one interview. There were five pupils on the Special Educational Needs ('SEN') register. The three of them, all boys, 'Charlie', 'Aaron' and 'Alan', were involved in the research.

There was one pupil from a religious or ethnic minority in the class. He was Muslim, being North African by birth. He was also involved in the research: 'Ahmed'. The remainder of the class were, like the class in the pilot study, a mix of degrees of religious commitment, though broadly culturally Christian. Like the pilot class, this class was also fairly representative of the school. The pupils therefore, like the pilot class, had little regular practical experience of ethnic or religious diversity.

6.3.2 The main research: the action steps

Did the research lessons go according to plan? Overall, the research process was maintained throughout the year, but a setback in terms of the teaching was my own absence, due to ill health. This meant setting cover lessons, which were not always followed exactly. The pupils' learning was therefore disrupted, and it

was not possible to apply as many of the techniques as frequently as intended.

The use of a *rapporteur* did not go successfully. As described in section 5.4.1, this is often part of a lesson plenary, when one pupil comes to the front of the class to give their view on whether the lesson aim has been met (Learning how to Learn 2002). After four attempts, I abandoned it and simply asked the pupils to review the lesson. This was because when one pupil was called to the front of the class the others would start to pack their books and equipment away; they saw it as a signal for the end of the lesson. As a teacher, I could insist that they got their books out, but this only took more time, and although I could extend the lesson into break, this would be seen as punitive, and would not be conducive to formative discussions about learning outcomes.

6.3.3 The main research: data collection

If there were some problems with the teaching, did the process of data collection go smoothly? The use of questionnaires proved unsuccessful. Obtaining the first set of questionnaires in October was unproblematic. However, the second set, at Christmas, was less so. A number were apparently being compiled by Harry Potter and his friends. Pupils had also been intrigued by the use of false names, and many were using them. As there was the risk that this data was becoming unreliable, I decided not to continue to collect it. Further, once the first set of interviews was complete, I had a clearer sense of the issues for subsequent interviews. Their purpose was to give an alternative perspective on the interviews, but as they were unreliable, there seemed little point in continuing with them.

By contrast, the interviews broadly followed the plan. The pupils were generally

interviewed in four groups of four, with interviews at the end of each term, giving twelve in total. The groups were chosen to ensure a mix of attainment, gender and attitudes in each. They were not grouped on the basis of their general educational attainment, or their attainment in religious education. The selection was based on my experience of them, after two months' teaching. While I wanted some members of each group who would stimulate the interviews, I did not want more introvert pupils to feel intimidated (Altrichter 1993, 103). The arrangements for the interviews are set out in appendix 9.

There were two problems with the interviews. Firstly, in two interviews, not all the pupils were available. Of the two exceptions, one pupil was absent from an interview in the spring term, and no substitute was possible, so only three pupils were present. Further, in the final set of interviews, another pupil was unavailable but a substitute was found. Secondly, due to my absences, the interviews were not evenly timed; for instance group D's interviews were in December, May and July, i.e. five and two months apart. However, neither problem is critical to the findings; there was only one new pupil, and the interviews were not intended to map progression over the year.

Is this data valid and reliable? This has been discussed above in terms of the research process and the data analysis (see section 5.3). As far as the accuracy of the transcripts is concerned, only one minor typographical change was needed. In general, Sikes' view is adopted that unless the statements are clearly not genuine, they should be considered authentic (Sikes 2000, see also Hookway 2003). There are however two important issues: firstly, any possible influence that I might have had over pupils. Secondly, however, I had decided to hold group interviews in order to mitigate this risk. But did this succeed, or create new problems?

6.4 Teacher researcher

Interviewing by a teacher researcher raises both general issues of research methodology, such as the problem of leading questions, and issues of the nature of the teacher-pupil relationship, e.g. one's authority. On the one hand, it may be that the power relationship severely damages the ability to collect valid or reliable data; thus Cohen, Manion and Morrison suggest that one of the difficulties of group interviews is how to 'get the children's teacher away from the children' (2000, 287). On the other hand, it may be that it actually allows pupils to speak freely and openly with someone whom they respect and trust: a teacher researcher could gather richer data than an academic researcher. Two examples of possible effects and two examples of openness by pupils are explored.

6.4.1 Influencing pupils

At times, I made leading comments or questions, which could have affected the pupils' responses. For instance,

> NF: So, the first question I asked on that was whether you thought you were getting better at RS generally. Um, a lot of people had ticked that, they thought they were, um I don't know whether that's true, or some people, they either agreed a lot, or they agreed. I don't think anyone disagreed particularly. Um do you think you've improved?
>
> Ellie: yeah
>
> Mike: yeah

Jake: yeah [B2[1]]

These pupils might have been unlikely to say that they had not improved. In the light of the relative paucity of such leading questions – only eighteen examples were detected - it is unlikely to be significant.

In a handful of examples, I could become over-defensive about aspects of the techniques trialled. In this example, a pupil considered that the use of module review sheets was fairly standard in many of his lessons, and I picked on a small design change that I had considered important:

NF: there's a column to fill in

David: they're slightly common

NF: common everywhere? Are they common everywhere with three columns? [A1]

The exchange ends there as other pupils continued the discussion. It is possible that my approach meant that he was unwilling to say anything further. This is a more serious misuse of the position of teacher-researcher; I was reacting because it was *my* worksheet that was under question. This confused my position as a teacher with the role of impartial researcher in a negative way; the examples however are very few.

[1] The reference indicates which interview group (A-D) and which term (1-3) from which the extract comes. It follows on all subsequent extracts. Extracts from the pilot study interviews are labelled PA to PD according to the four groups.

6.4.2 Reactivity

Reactivity is term for the problem that participation in research can alter the ways that participants behave or respond (Kumar 1995, 97). Here the point is not that I might influence the pupils but, because of my position, they might try to influence me. For instance, in a discussion on the content of lessons,

> Alan: because RE shouldn't be like a written subject, it should be like a mental subject
>
> David: yeah
>
> NF: just in your head? And I've made you do lots of writing have I?
>
> Alan: yup
>
> *Laughter*
>
> Alan: hint, hint [C1]

Alan was openly trying to influence the nature of future lessons – no more writing. While the pupils are not lying, this material obviously therefore needs to be treated cautiously. More positively, it at least suggests that pupils were not intimidated by my position.

6.4.3 'Teacherly' reactions

Questions or issues sometimes arose that I dealt with as a teacher would. This could include pupils' behaviour in the interview, e.g. by fiddling with the recording equipment, and could also include issues of content. For instance in a discussion about lessons on religious views of evil:

> Kevin: if God was perfect, then there would be no evil, he would –

Ahmed: well, he's allowing us to see what you would do

NF: the classic, the classic defence is that it's more perfect that we have free will, and we have the ability to choose. God could have created us to be perfectly good all the time

Annie: yeah, but

NF: but, if he, but if we weren't, we were able, if we were programmed to be good all the time, then we would not really be good. Goodness is only good if we can freely choose it. And to freely choose it, you have to be able to freely choose the alternative, which has to be [possible] [D2]

Responding like this would obviously re-assert my status as their teacher, by dominating 'classroom' talk. However, it would have been awkward not to deal with the issue as it had arisen. There is no evidence that the pupils were confused by this change of style. Indeed, one could argue that it encouraged pupils to open up more fully; by discussing the lesson content rather than the self-assessment techniques they were able to judge their progress more effectively in the interview.

6.4.4 Pupils' exasperation

A more complex issue was when pupils expressed frustration with my interviewing approach.

Kevin: I don't like this, everything you say, or everything <u>we</u> say

Annie: gets turned around

NF: *laughs*

Kevin: <u>you turn it around</u> so it makes us look, look better. No, you just turn it round somehow and it makes, I don't know [D2]

This example shows that the interview process did frustrate pupils at times. Perhaps this was because they were encouraged to look more deeply than they wanted to. Kevin felt that I was transforming his words in a way that misrepresented what he had said, even if I made it 'look better'. However, it also shows that pupils were not afraid to express their frustrations; even if they felt that I was misrepresenting them or asking too much, they were prepared to challenge this. Further, the benefits of a group interview can be seen; other pupils support the challenges, for instance Annie supports Kevin, which might not have been possible in individual interviews.

6.4.5 Teacher researcher: conclusion

Any teacher researcher needs to be reflexive about their dual role: in the transcripts, the tensions between them are apparent. I could behave in 'teacherly' ways that were likely to influence the pupils, by asking leading questions, by dominating discussions, by misrepresenting pupils' views to them. There is also evidence that pupils wanted to influence my teaching through the research; it was their opportunity to give feedback. Nevertheless, while this can only be decided on a case-by-case basis, it would seem that overall there is unlikely to have been a significant amount of influence, and it seems more likely that the situation, generally, allowed pupils to speak openly and freely. At this micro-level, the practices of teaching and research were generally combined successfully.

6.5 Pupil relationships

Having group interviews may have helped solve the issues of my influence. However it created another set of problems in the pupils' influence on each other.

Their own relationships came into play: did this affect the trustworthiness of the data? Generally, the pupils seem to have responded honestly, but there are three problematic issues. By contrast, there is an example of a considerably more positive consequence, as pupils became co-interviewers.

6.5.1 Social activity

One problem was that the interviews were often a chance for pupils to chat, joke, argue or banter with or about their classmates (see Cohen, Manion and Morrison 2000, 287). For instance, one pupil made fun of another pupil's response to a question:

> David: the main thing is you can walk out of the classroom in the morning and say to yourself, I've actually achieved something this lesson. [NF: OK] the aim is [*laughter*] the aim is [*laughter*] no the aim is, right
>
> Charlie: the way David went, right [*laughter*] "walk out, morning, he's come out you've actually achieved something this morning" [A1]

There is a tension here. The group interview helped ensure that my influence was not too great, but meant that there was peer pressure within the interview itself. It was in these situations that a degree of behaviour management was called for, and so led back to the teacher-researcher problem described above.

6.5.2 Games

What also emerged in one interview was that the pupils set themselves a challenge. We were discussing whether the transcripts needed my questions as well as their comments:

NF: well the trouble is you need the context because otherwise if you haven't got the questions,

Jenni: but then we'll know how brainy your words are, because some of them are pretty like

Sara: like 'technically'

Jenni: don't take the Mickey

Sara: aren't you Jenni? You are going to try to put that in five times [C2]

For these pupils their own private game was to be part of the process. This does raise issues of reliability. Are they more concerned with finding an opportunity to use the word, or to express their genuine thoughts? This again is a case by case decision. On the whole, whenever one of them did use it, the others would comment, so its effect is obvious. Furthermore, they were prepared to admit to the game, and the word itself is unlikely to compromise the general tenor of their remarks.

6.5.3 Evolving collective thought

Thirdly, there was also the possibility of pupils' ideas evolving through the interview discussions, rather than being a personal response. At times, it is not clear if pupils are saying what they think, or simply helping flesh out someone else's answer, and imitating any changes to the other's ideas. Consider Billy's and Kevin's responses on the module review process:

NF:...What about the, the tick sheets, the module review sheets?

Kevin: I found them boring for some reason

Billy: you don't pay so much attention so you tick the box you want really.

NF: OK

Kevin: 'yup, oh that one, yup' if you do it properly, if you do it right,

Billy; yeah it is pretty cool

Kevin: yeah, it is pretty cool, because at the end you think, 'oh I know what that is now', because you think back when you was doing it and you was thinking 'what the hell's that?' [D3]

At the start, Kevin finds the process boring, and so Billy explains how he does not take the process seriously. Then Kevin seems first to describe a casual approach to box-ticking, but then focuses on 'doing it right'; Billy then agrees with the idea of doing it right, and then Kevin seems to convince himself that it is a useful process, as it moves from 'boring' to 'pretty cool'. They leapfrog over each other to get to the opposite point of view from where they started: do they like the techniques or not? This raises an issue with interview technique, and whether I should have probed more deeply.

5.5.4 Pupils as co-interviewers

The positive example is when pupils took control of the interview themselves: there were twenty-three examples of this. Sometimes, they pointed out a need to return to the content of the interview, rather than an irrelevant topic, e.g. Jake commenting 'well this is wildly off topic' [B2]. They also interrogated each other when they considered that the answer was inadequate; this example is about work on the holocaust.

NF: right, but in terms of that particular thing you didn't like it because...

Annie: don't know, I just didn't like it...

Ahmed: you <u>must</u> have a reason – you <u>have</u> to have a reason [D2]

They would also check to find out if other pupils were paying attention to the interview:

> NF: OK, so that's a kind of practical benefit. Annie what about you? Do you agree with that, disagree with that?
>
> Annie: yeah
>
> Kevin: what was I saying?
>
> Annie: talking about putting your opinion across at home [D3]

They also gave others instructions when they felt that they had not been participating:

> David: you don't know what we are talking about
>
> Helen: *nervous laugh*
>
> David: Helen, you need to say something, at least a paragraph, now. [A3]

Finally, they would conduct their own discussions about some aspects of learning, e.g. a visiting speaker:

> Sam: I think the holocaust stuff was good as well because - it just was, I think. It was interesting to find out about it and what the person who came in, he survived, he survived it as well. It was pretty good.
>
> Aaron: I think that evil one was =
>
> = Alan: not really, that wasn't very good at all.
>
> Aaron: wasn't it? I thought > Sara: no because <
>
> Alan: he didn't really survive the holocaust. He wasn't like in a concentration camp. He was just there, at the time.
>
> Aaron: in the right place.

Jenni: I don't know I wasn't really listening

Alan: umm

Aaron: I thought it was really good. I thought it got an inside view instead of like going through the textbook and stuff really. It was really good.

Jenni: I'd much rather saw it on video because then you can like stop it. [C3]

In this example, the interview almost becomes a focus group instead, with the participants 'discussing a topic supplied by the researcher' and interacting with each other (Cohen, Manion and Morrison 2000, 288). Focus groups have their own strengths and weaknesses (Morgan 1988); however what is clear is that these pupils are not intimidated by being in a group.

6.5.5 Pupil relationships: conclusion

In this section, three problems have been identified: interviews as social activity; pupils' games; evolving collective thought. Nevertheless, the final section shows clearly that the pupils were generally engaged in the research process and broadly saw it as valuable. Thus they were the ones who overcame some of the difficulties with group interviews, such as avoiding the researcher being the authority figure, keeping the interview relevant, and eliciting genuine responses from the other pupils (e.g. Cohen, Manion and Morrison 2000, 287). While of course each particular exchange needs to be considered individually, group interviews seemed in general to be a fruitful way of getting at pupils' genuine views.

Overall, this shows that the interview data is generally valid and reliable. I did not influence the pupils as their teacher. In fact, it can be argued that the pupils were more open to the process because I was a familiar figure, and because of the

use of group interviews. The combination of these two elements created a new set of problems – the management of the group. However, the last example shows that the pupils understood the structure of the research process; they felt able to ask questions of each other, and indeed to comment on each other's contributions. Lewis asks that:

> Teacher researchers consider carefully whether or not the children's inter-relationships or his or her relationship with them distorted responses. (Lewis 1992, 417)

I consider that on balance they did not. More than that however, it can be argued that this is particularly rich *because* this is teacher research. The relationships in the interview grew out of the classroom relationships, particularly in a subject where debate and discussion are prominent. Moreover, as in classroom discussion, pupils could challenge and develop each other's ideas. The interviews are rooted in discursive lessons, and are a reflection on them because the topics for discussion were shared with the pupils and teacher researcher. Indeed, data collected from within a relationship of trust and respect may be more fruitful than data collected by an academic researcher.

6.6 Objectivity in data analysis and theory generation

The next stage in the research was the analysis of the data. This was carried out using a mix of deductive and inductive codes, as discussed in section 5.4.3 above. In all, there were sixty-nine codes, hierarchically organised into nine categories (following Rubin and Rubin 1995, see appendix 8). There was however the danger that I would simply read into the data what I wanted to find there. An independent educational researcher coded a sample transcript using

the codes I had created, to ensure that the coding concepts were conceptually valid and had category reliability (Asher 1984, 15). This gave an inter-rater reliability coefficient of 0.9 (number of agreed coded sections/total number of coded sections), which is higher than the acceptable level of 0.8 (Herzog 1996, 43). Only one code was ambiguous (6aiii, on pupils' views of their learning across school), and clarification of the theme of the code resulted in greater clarity on re-coding. These codes therefore are conceptually coherent and consistently applied, though doubtless other researchers could interpret the material differently (see e.g. Silverman 1993, 47). Some of the material was not directly relevant, notably the material on whether teachers should express their own beliefs in the classroom. I have discussed it elsewhere (Fancourt 2007), and refer to it where appropriate.

6.7 Conclusion

This chapter underpins the quality of the research by transparently setting out the practicalities of conducting the teaching and the research. It has shown how reviewing the pilot study affected the main research design, notably in shifting from questions about effects to questions about experiences. The reasons for abandoning one of the action steps (the rapporteur) and one of the methods of data collection (the questionnaires) are explained.

Secondly, in terms of the validity of the data collected, the strengths and weaknesses of the interviews have been considered rigorously. This included both the issues associated with a teacher researcher: the dangers of influencing pupils, reactivity, 'teacherly' reactions and exasperation. In particular, the strengths of interviewing by a teacher researcher are highlighted in an distinctive way. It also included group interview issues; the interviews as a social activity,

games by pupils, co-interviewing and the evolution of collective thought. Generally, the data has been shown to be valid. The coding process and theory generation have also been scrutinised, and found to be dependable. It is at this point therefore that the findings from both the pilot and main research can be addressed with confidence.

Chapter Seven

From techniques to themes: analyzing the pilot study

In this chapter, the main themes that emerged from the pilot study are set out. Sometimes, 'exploratory research' is used to designate preliminary research to identify possible issues, whereas a 'pilot study' is a feasibility study, in terms of the methods (e.g. Kumar 2005, 10); here the term 'pilot study' is used to cover both feasibility and exploration. The exploration is as important because it helps to map out the landscape of pupils' experiences of assessment. Furthermore, feasibility and exploration are entwined because the range and depth of the emergent issues reflects on the main research design, particularly in terms of the focus of the research questions and the management of the coding. The two are inter-related.

First, the pilot research problem and research questions are explained, then the questionnaire data and interview coding described. Pupils' views of their progress and their general understanding of self-assessment are set out before dealing with answers to the four research questions individually. Four other generic themes are explored, before reviewing the findings from the pilot.

There were two differences in method between the pilot study and the main research. First, as explained above, I used questionnaires in the pilot study, but abandoned them in the main research (see section 6.3.3). The questionnaire data is included here, as it sets the scene for pilot study interviews. Secondly, since the codes were partly inductive, the shape that the coding and clustering took differed from the main research; it also became considerably more refined.

Ideally perhaps the pilot data could be re-analysed more insightfully in the light of the more detailed main research codes; however for transparency's sake it must be judged on its own merits.

7.1 The pilot research questions

The research aim and questions for the pilot study were (see also section 6.2):

- Which self-assessment techniques are most suitable for the different aspects of learning in religious education?

  1. Does traffic-lighting provide help with any aspects of learning?
  2. Does the use of a review in each lesson help with any aspects of learning?
  3. Does the formative use of summative assessment criteria help students develop their evaluative skills?
  4. Does the development of class assessment criteria help the development of empathy?

The aim focuses on the suitability and the questions on the helpfulness of the techniques. This structure of a series of questions on individual techniques shaped the questionnaires and the interviews.

7.2 The pilot questionnaires and interview coding

The results of the two questionnaires are set out in appendices 10 and 11. The questionnaires were carried out in February and April 2004, and there were some differences in their design. The second questionnaire, but not the first,

was analysed by gender in order to consider whether this played a part in pupils' responses, and it also included a final question asking pupils to rank the various self-assessment techniques in order. The issue of gender was not pursued in the main research; it raised many additional theoretical issues (e.g. Gipps and Murphy 1994, Weiner, Arnot and Davis 1997). The data from the questionnaires is given in percentages.

The interviews were coded into fourteen categories, which are set out in appendix 7. The first eight arose out of the research questions and were thus deductive; the next six were new ones that arose from the data itself, and were inductive. A combination of approaches is seen by Altrichter *et al.* (1993, 124) as methodologically acceptable; the issue is always sensitivity to the respondents' meanings. An approximate indicator of the proportion of pupils involved is given, in order to help the reader get a flavour of the response, not to suggest greater validity or reliability.

7.3 Pupils' sense of progress

Since the pilot research aim and questions focus on improvement, it is important to know if the pupils thought that they had improved; clearly the techniques could not help them otherwise. On the first questionnaire, questions 1 to 3 showed that all pupils considered that they had understood different explanations of evil, most had evaluated these explanations, and almost all had given their views (appendix 10). The second questionnaire gave a similar result for this and for the development of empathy (question 2), though less for the development of tolerance (appendix 11, see section 7.5.4). All pupils therefore considered that they had achieved almost all educational goals.

7.4 Pupils' understanding of self-assessment

Did the pupils have an understanding of the process of self assessment? About a half of the pupils could articulate a clear account of it:

> Bill: You know what to do next time round um because you can see the areas that you need to work on and improve and areas that you're, you're already working at a high level at, and you can see how well your work is going. [PB]

> Tim: It will help you go over your work, and like look at the bits you did understand and look at the bits you weren't too clear on and stuff. [PA]

By contrast, many others were not so clear.

> Kath: I don't really get what to do with self-assessment; we do the sheets, and things.
> NF: what you mean? The tick box sheets and things?
> Kath: I don't really get what to do with it, and things, and I don't really like them, I think [PB]

The questionnaire data was similar. The first questionnaire showed generally a positive attitude to it (question 4), with 66% in favour. The second questionnaire gave a less positive response to this question (question 4), with, overall, 48% in favour. However, neither of these figures bears much relationship to the results of the questions about individual techniques, which I

intended as a way of breaking down the more general question. Thus on the second questionnaire, while 54% thought that self-assessment helped them learn, 88% thought that levels helped them assess their work. If the pupils interpreted it as I had intended, then the remaining 34% who liked using levels ought also to like self-assessment generally. This suggests ambiguity in the wording of the questionnaires, and a lack of construct validity.

7.5 Answering the pilot research questions

With this in mind, it is appropriate to turn to the pilot research questions, since the pupils seem to be clearer on these. What did they think of the individual techniques?

7.5.1 Traffic-lighting

On the whole, pupils found this helpful, but this comes with some qualifications. The first questionnaire, which only asked about the lesson review process showed 57% in favour. For the second questionnaire, which asked about both lesson review and module reviews, their views were more favourable, with 73% in favour. The second questionnaire, question 5, shows boys slightly more in favour than girls (76% to 70%), but in particular, the boys were more enthusiastic, with four agreeing strongly (30% of the boys), and one girl disagreeing strongly. The last section of this questionnaire shows a slightly different story: 31% of the girls, but no boys, chose it as their favourite form of self-assessment. Overall, taking all preferences into account, it was one of the least favourites for boys, but the second favourite for girls; for both it emerged as the second favourite technique.

The interviews showed a similar pattern of mixed views for the various aspects to traffic-lighting. There was strong approval for the lesson review, for instance:

> Claire: I think it helps you like clarify what you've achieved, so you can like just make sure that you've achieve what you're meant to. I think it's quite beneficial for that reason. [PA]

On the other hand, about a quarter were cautious or negative,

> Tim: it helped when you looked back through it, if like you need to clear up points on it, but besides that it's not like you know you need it, but it does help when you look back through it, I think. [PA]

The module review process met with similar comments of approval.

> Ruth: I think it helps because um I know what I done and learnt and can look back (**) um think what it is about um I can look back on it and just like say um "I've done this" or "oh yeah, I can remember that" 'cause if I know I didn't do it then I would never remember it and, you know, just like forget it totally, so I find looking at the tick boxes that I've done it then I think it's the best. [PC]

There were two main qualifications. First, pupils found it difficult to decide if they 'knew about' or 'could do' something because they did not know what they should be comparing it to.

> Charlie: yeah they worked to a certain extent...sometimes you don't

know if you are right in what you are saying, that's the problem. [PD]

David: The thing is, how do you know what you don't know? [PD]

The second reason was because some pupils thought it was meaningless:

NF: so which ones for you were a waste of time?

David: um, well, the traffic-lighting one really. Where you like have the titles and you have got to decide whether you know it or not; 'cause you're going to get taught it anyway. [PD]

David's remark shows how the mere techniques of assessment for learning may not be enough for pupils. He wants a more responsive approach to assessment and teaching as a whole, not just the opportunity to identify his own progress (see Deakin Crick *et al.* 2007). For more autonomous learners, what a teacher thinks is the adventurous use of formative assessment might not be enough.

Finally, though a couple simply did not engage with it; it was another school task they were asked to carry out:

Sara: It's just kind of like a sheet you just go through and then you kind of just forget it...[PA]

This points to the issue of getting students to engage with a new approach to learning; they are content to 'get by', even with processes that should liberate their learning (Perrenoud 1991, 92).

7.5.2 Formative use of summative assessment

It is argued that a reason why students are concerned to 'get by' is that they have become too 'ego-focused' as a result of summative assessment processes. Was it possible for pupils to interpret marks or levels in a formative way? Butler had suggested not (Butler 1997, 1998). Others are more positive (Harlen 2006). Many of the pupils liked to have levels:

> Helena: was the best? Using levels.
>
> Charlie: *grunt of approval*
>
> David: I liked using levels [PD]

This however could simply indicate an ego-focused response. One pupil described how he would use level descriptors formatively:

> Bill: I read through mine when I was writing the work. I put down whatever came into my mind and then, like Jack did, I didn't really read it after I had done it, just like looked at the odd paragraph or words or sentences, but um when I got my book back and I looked through it I thought "actually this is quite a good piece" and I changed the level that I put for myself. Then I changed it to, um, I think I put a 4 for it before and I put it to a 5, but I really got a 6 for it and I thought well maybe I should have read through it and actually saw if I could make it better, even though it was better than I thought it was going to be. [PB]

This was also borne out in the questionnaires. In the first questionnaire, questions 6 to 11 dealt with a range of issues linked to this. What was

particularly strong was the use of peer assessment techniques to understand the criteria for their work: six agreed strongly, fourteen agreed, and only one disagreed, i.e. 95% were positive about it. Furthermore, 95% were positive about the feed-forward aspect of the process (question 9), with nine pupils strongly agreeing. The second questionnaire bore this out; question 6 shows 84% agreed. The boys were very enthusiastic – all were positive, and 54% agreed strongly. The girls were less enthusiastic. The final section however shows the strength of feeling for the use of levels. It is clearly the most popular first choice of technique for both boys, girls and altogether

This suggests that Butler's research is not as generalisable as one might assume. This may be because of the format within which marks or levels were given. In her research, it was individual feedback after a piece of work. My research was by using peer assessment to grasp the criteria involved. This may account for a more positive response, especially in a wider school setting in which peer assessment is used in this way, which would reduce pupils' antagonism to marked feedback, as Maqsud and Pillai suggest (Maqsud and Pillai 1991).

7.5.3 Class assessment criteria

This technique was carried out after the first questionnaire, and the first interview. The second questionnaire shows the pupils' views of it (question 7 and final section). 65% were positive about it, while 35% were not. There was no real difference between boys and girls. In terms of choice of technique, only two pupils, a boy and a girl, chose it as their favourite, but interestingly these were the two 'gifted and talented' pupils. It emerged as the least favourite for

boys, girls and altogether.

In the interviews, about half the pupils were positive about the process:

> Charlie: yeah, because um you know yourself better than anyone else and you know what you need to, rather than being told what you have to do, you can actually sort of set yourself a target and stuff for what you have to work on if you know what you need to do. [PD]

The lesson itself was animated at times. One group had a long discussion about whether you would notice the sort of gun that a member of the SS was using when coming to arrest your family: was this good use of detail or in reality would you simply not notice this sort of thing? The nature of quality was itself contested.

However, much of the interview focussed on the fact that it was based on peer assessment of exemplar work before they did it, rather than the precise issue of class criteria:

> Kath: I'd say it was better than using the levels, because if you look at pieces, other people's pieces of work and then say what you think, if it's good, then you're going to put it in your work. [PB]

As a teacher, I would say that the overall quality of the work was better. The pupils used a variety of the features they had identified effectively, such as a sense of narrative, and relationships, e.g. between child and parent. However, one issue for some was the question of originality. They were concerned that

taking other people's ideas, even simply as criteria, was not original, and therefore was a form of copying.

One concern with this process is the relationship between self-assessment and peer assessment. While peer assessment is a vital way of helping pupils understand the criteria, and of breaking down concerns with feedback, it is not quite the same thing as self-assessment. It certainly helps pupils, like Jack, to include criteria in their work, but this arguably is more a process of grasping quality, and then realising it, than self-assessment.

7.5.4 Discussing tolerance

The discussion of tolerance as a result of studying Jewish reflections on the holocaust was carried out very late in the term, and therefore only the last interview (Group D) discussed it. Pupils were asked to write a sentence and explaining if they felt learning about the holocaust had made them more tolerant (see appendix 2, for an example). My diary has the following entry:

> Brian made the point that it helped if you knew what they were feeling. Jon said it helped you to get things in perspective – that your troubles were not quite so bad. 3 or 4 agreed with these views, especially Jon's. Some (e.g. Josh) felt that they were tolerant anyway, and this was useful, but did not change them.... Tim thought that it made him sorry for those who went through it, but not for Jews generally; he compared it with being sorry for a whole tutor group because one person in it had experienced suffering. Jamie thought that it was only useful if you were going to experience the same things.

> We then discussed if it was possible to empathise completely with the escapees on the *Kindertransport*. All [who spoke felt] that you could not imagine it fully. But Josh and Jon felt that it was still valuable to try. Claire connected it with making sure that it did not happen again.

The results of the second questionnaire reflect some of these difficulties. Question 3 shows whether the pupils felt that they had become more tolerant as a result of studying the module. Overall, 65% thought they were. But the interesting difference was between boys and girls: 77% of the boys, with 23% very positive, against 54% of the girls, with none very positive, and one very negative. The boys' enthusiasm is borne out in the result of question 9 too. In this, overall, 62% thought that having the discussion helped them assess their tolerance, with 92% of the boys being positive, against only 31% of the girls. Indeed, on the whole the girls found it unhelpful. This trend is replicated in the results of the last section. It was the second most popular first choice for boys, but one of the lowest for girls, and overall it was the second choice for boys and the third for girls.

In the fourth interview, there were three boys and one girl. Two boys were enthusiastic:

> David: I liked it...'cause um when I first saw that question on the board I couldn't really make my mind up, but then after I'd heard a few other people's ideas it helped me to realise what, whether I thought I was more tolerant or not.
>
> Charlie: when I first saw the question I didn't really know what it meant and particularly by listening to other people I sort of started to know what

> it meant and started to make up my own mind...about things. [PD]

The other boy did not think his values had changed:

> Brian: um, yeah I like...the talk didn't help because I knew how tolerant I was, and I know what it meant, but it wouldn't change me really...because I had really like made up my mind. [PD]

For the girl, her concern was with the group size, not the process of discussion itself.

> Ruth: um I didn't really like big discussions of like everyone 'cause I don't like sharing my views with everyone, so I would have only done it with friends. [PD]

Discussion, however, was considered to be the only way to reflect on this issue:

> NF: OK. Do you think for tolerance we could have used any of the other techniques...?
> Charlie: No. I think the talk was better for tolerance because you could relate to everyone else's point of view. [PD]

For Charlie, it was a useful way into a difficult process of self-assessment, because it was only though discussion that he could grasp what the concept was.

One important issue was whether they were now more tolerant. Some, e.g.

Brian, felt that they were tolerant already, and therefore they had not changed in this. The definition of tolerance, however is, is contested (e.g. Afdal 2007, 86-133). Discussion seemed to allow pupils to start to develop and apply their own understanding of tolerance.

7.6 Theoretical implications

The specific nature of the research questions in terms of improvement did not allow for some of the other themes to emerge. It was also apparent that there were four more general themes: differentiation in assessment for learning; the use of levels; assessment as a social activity; sharing criteria.

7.6.1 Differentiation and variety

Different pupils are better or worse at different self-assessment techniques, and further, they have different preferences. Many like the use of levels, but others really like discussion; some like using traffic-lighting, but others prefer devising their own criteria. This is obvious when one thinks about it; just as pupils have aptitudes and preferences within a subject, so they also have skills and aptitudes within self-assessment. For instance,

> David: um I think that some self-assessment ideas work, but other don't, um not really... just a bit of a waste of time.
>
> NF: so which ones for you were a waste of time?
>
> David: um, well, the traffic-lighting one really. Where you like have the titles and you have got to decide whether you know it or not; coz you're going to get taught it anyway. [PD]

Brian: um...some...I don't like really doing self-assessment, like marking the work but just in general, like traffic-lighting I do like that...[PD]

One judgement that a teacher would have to make is about how 'good' a pupil is at self-assessment. It bears out Laveault's suggestion that there are three pedagogical principles: self-assessment must be diversified; it must be differentiated; it must become the object of progressive training (see above section 2.4.5, Laveault 2007, 233, see also Kincaid 1991, 123). An essential part of a teacher's professionalism would be how to develop and maintain task-focused pupils. The changes in the teachers' expectations of their students would mean that students had to change their role in the classroom. In particular, they would have to change from passive recipients of the knowledge offered by the teacher to active learners who were able to, and were expected to, take responsibility for and manage their own learning (Black *et al.* 2003, 97). Where is a particular student on this spectrum? And what techniques are appropriate at each stage, in each subject? Should one abandon any formative use of summative assessment until the pupils are clearly task-focused, and then introduce it, or should it be a continual part of the process?

This process of differentiation of assessment would have to be combined with differentiation for different abilities. But what would be the effect of giving the more task-focused pupils, whether high or low attainers, their marks or levels because they can interpret them formatively, but not to give the ego-focused pupils their marks or levels? Some pupils recognised the need for a range of techniques:

Tim: I think it would be good to see, like um, a different way of looking at

it, but I reckon marking would be better, but it would be good to see how it works, it's the whole class and that lot. [PA]

This shows both an understanding of the usefulness of variety, even though he had a clear idea about what he considered to be the better approach, and recognition of the shared nature of the learning process.

7.6.2 The use of levels

What follows from this discussion, and is central to the debate about the value of assessment for learning as against more summative approaches, is the use of grades or levels in feedback (see section 2.3.2). What light is shed on this issue? The overwhelming preference on the questionnaires for using levels was surprising. Clearly, the pupils, both boys and girls, thought that this was useful for them. On the first questionnaire, I asked two questions (10 and 11) as to whether the use of levels motivated the pupils. Twelve *strongly* agreed that they liked to know and nine agreed; none disagreed at all. But three pupils thought that knowing their levels did not motivate them, even if they liked to know the levels. They are not however therefore saying that knowing their levels is de-motivating; it may be motivationally neutral, in that they are motivated or de-motivated by other factors.

The interviews add more insights. The levels were important to many interviewees, in different ways. Some pupils made rather ego-focused remarks:

Laura: I don't really know what you have to do for the levels but I know which levels are good and which are bad but er...[PA]

Sara: I just like looking at the level at the end. [PA]

A few revealed a more task-focused approach:

Claire: Comments. I find that helps me more. It like enables me to see where I am at and like whether I understand it and it tells me what I got to do to improve, or like whether it's good or not and like it really helps me more than a level. I'm not that bothered about the level, really. It's just whether the teacher thinks it's good or not, and their comment. [PA]

Others occupied the middle ground, wanting levels, but recognising that formative comments were useful as well:

Tim: because you want to know what they think of your work and if they're just going to tick it what's the point? I knew that was right, but I want to see what you think of it.
NF: so if I was to just put a level...if all you're going to look at it the level, why can't I just put a level?
Tim: No, I'd read the comments, but I only look at the mark first, because it's like "what did you get?", "I got..." and then you're like comparing it, but I did look at the comment afterwards. [PA]

Charlie: yeah because you become too dependent on the levels.
NF: why would it?...
Charlie: well, I think you're sort of, if you were given, well just grade yourself by levels and you're sort of given nothing else you'd sort of be lost. [PD]

One aspect of the pupils' position on the ego-focused to task-focused spectrum (Gipps 1994) was the use of levels in self-assessment. There was a range of views:

> Tim: It's good seeing like what like teachers have written down and comparing your level to theirs but it like, it's like (*hard?) what do I put down?...I don't like marking my own work. [PA]

> Brian: oh, no I liked using levels...a lot because it gives you more of an idea of where you are...You know what the like average level is, and I think it helps... you know where you are, below that or above it, or on target. [PD]

Tim's view is task-focused in terms of his desire to compare his view of his work with the teachers. He is clearly in control of the learning, but at the same time he does not like giving his work a mark or level. Brian likes using levels in this way, but to compare his level with a notional average, in a rather norm-referenced way.

The question is whether pupils can appropriate the criteria for their own educational ends, and what the teacher's role is in this process. In particular, it is the use of grades or levels for 'ipsative' referencing (Black 1998, 73, see section 2.5.4). My research suggests that most pupils in the class were able to 'decode' the levels into meaningful information for improvement, even if many were also treating the information in a norm-referenced way. This of course is in the context of a mix of techniques, both summative and formative, and not just this one technique.

If there is a sense of *quality* of work, then pupils need to be able to share it, and crude though they are it seems that graded criteria are hard for pupils to manage without. Arguably, it is more helpful to use them transparently in this way if they are part of the educational landscape. But this is not the same as assessing summatively, and then trying to use the data formatively. There is an issue about the primary motive of the process. If it is to be formative, then the formative function, using ipsative referencing, needs to take priority. The summative and formative functions cannot be equal. An ipsative approach requires the levels to be seen by all in the classroom as a simple way to identify certain qualities in pupils' work, which can be developed and improved on, and not as cast-iron descriptors of ability. The problem is in ensuring that the possible negative effects are minimised.

7.6.3 Assessment as a social activity

Many of techniques described above require a pupil to assess his or her own and others' work during the lesson. This means that they will see each other's work, and has social implications:

> Sara: you are worried that you are going to under-mark yourself which means people are going to think you are too modest, or you are going to over-mark yourself which means that people are going to think you are big headed. [PA]

It contains two dangers. First, if pupils get given their levels or marks summatively there will be social implications in terms of the comparison with other pupils. Indeed one reason why summative assessment is considered de-

motivating is because pupils compare results (Gipps 1994, 41). This will always be a problem with the use of levels in some way, even if used ipsatively, as described above. One pupil outlined the practical problems.

> Ruth: well, I do like knowing what level I'm at, because I am really interested in what level I'm at. I don't like the whole class knowing what it is, unless it's a good mark because if it's really bad and everyone has got eight's and I get a five or maybe a three then it doesn't look that good, but um I don't really want anyone to know. I mean everyone gets their grades and sometimes you ask them privately and then everyone asks you and you ask your friends and that's fine, because you don't mind them knowing. [PC]

This is particularly a problem in terms of relations between friends and the rest of the class as a whole.

> Helena: um, I didn't really like big discussions of like everyone 'cause I don't like sharing my views with everyone so I would have only done it with friends, so
> NF; you'd have preferred it if we'd done it with smaller groups of friends?
> Helena: yeah. [PD]

As in many classroom activities, there are issues in dealing with shy pupils, who do not like to talk in front of, or read to, the whole class.

Other pupils were happy for their work to be singled out in this way, and indeed can find it rewarding:

> NF: Rob, you were the one who had your work handed round...
>
> Rob: yeah, I didn't really mind 'cause it was quite good 'cause no one really had anything majorly bad to say about it, so it gave me a little bit more confidence. [PB]

However, one danger is that pupils can use the process to insult others. For example, during a lesson in which pupils had to peer assess work using levels, Kayleigh's group gave a low level to work by Greg. On this, other pupils criticised them: "'I bet that's better than what you got". I interpreted this as demotivating for the group because they were being judged for their failure to grasp the criteria. However it later emerged that this was possibly not the whole picture:

> Charlie: they don't really like Greg
>
> NF: sorry?
>
> Charlie: it's only because Kayleigh doesn't like Greg, actually. [PD]

Thus the process of giving levels was possibly another way of continuing a social dispute, and undermined the function of the exercise.

7.6.4 Sharing criteria

Finally, at the heart of this process is the issue of sharing criteria, and how pupils come to own them: 'enculturation' into the assessment community by pupils (Wiliam 1998). Pupils cannot self-assess without knowing standards by which progress is to be judged (see Sadler 1989, and Black and Wiliam 1998a, 30). The pupils identified problems in gaining this understanding.

David: um, I think it like helps you to like know what you've got to try and achieve in the lesson, and, but that's about it really. It like sets you on target but I'm not really sure if it like helps you to assess what you are doing, whether you did well in it or not. [PD]

Charlie: um yeah, I think (*the level-path sheet's) good but like I said before the problem is if you do it, you don't really know if you are right or not, so you might put it too high, or you might like put it too low. [PD]

The problem was neatly summarised:

David: The thing is, how do you know what you don't know?
Charlie: That's, that's a problem: if you've got a topic that you don't know much about at all you can't really select things [NF: OK] about it. [PD]

David was the 'gifted and talented' boy in the class, but he preferred setting class criteria to using levels (as did the 'gifted and talented' girl). He knew that it is hard for pupils to decide how well they are doing at something if they do not know what 'well' means. His comments are striking because he took the process very seriously, and understood fully what was involved.

On the other hand, what did emerge was the juxtaposition of self-assessment with peer assessment and teacher assessment. Pupils found it helpful to assess other pupils' work:

Helena: Yeah, marking other people's I think does
NF: right

Helena: if you look at really good work then you sort of know how good it is and how good yours has to be

NF: and that's useful? >Helena: yeah< Do you think you were then able to decide how good yours actually was?

Helena: No I think I don't like really marking my own work. [PD]

This gave pupils an indication of quality, as was shown in the empathy work, and in the use of levels in the feedback on the essay. Further of course the use of and reliance on teacher assessment is also mentioned above. The problem is in knowing how much teacher assessment is appropriate in order to enable pupils to carry out self-assessment. The teacher is after all the expert in the classroom, the person who has seen many previous examples of that completed task, and who can share that expertise.

7.7 Discussion of themes in the pilot study

This analysis clearly shows that issues with particular techniques in the religious education classroom cannot be separated from general issues in pupils' ideas of and reactions to assessment. The ways that techniques worked for students were bound up with their approach to assessment processes generally. The methodological implications of this are set out in section 6.2.4 above, notably restructuring the research questions, more detailed coding and the need to focus on achievement in religious education. However, it also has implications for an understanding of assessment in religious education because it shows that this will be coloured by pupils' more general experiences, in what Ecclestone and Pryor call their 'assessment careers' (see section 2.6, and Ecclestone and Pryor 2003), rather than by teachers' intentions. However, it is

only through trying new techniques that pupils may come to change their perception of both assessment and religious education. The new techniques could challenge pupils' preconceptions and thus modify them; their reactions are not cast in stone. The 'failure' of one technique may be part of this contestation, and the 'success' of another technique may be short-lived, as pupils adjust their approaches to learning.

7.8 Conclusion

The analysis of the pilot study in this chapter has pointed to significant and original issues for the research. It is significant because it addresses pupils' views of how they could develop the ability to self-assess. The pupils in the class thought that they had made progress in religious education, and explained which techniques they found helpful; in particular they highlighted the use of levels, which was surprising in the light of the literature (e.g. Butler 1987, 1988). However the pilot study also showed that pupils' reactions to the individual techniques varied. They could be seen in the light of more general issues to do with their experiences of and approaches to assessment (Ecclestone and Pryor 2003), such as how they dealt with levels and marks, the social interactions within the class and how they understood the criteria.

One can begin to see how the two fields of investigation, generic research on assessment for learning and issues of pedagogy in the classroom, are inter-related in practice, and this is also original because it appears to be the first data to demonstrate this inter-relationship. The literature review did not reveal any other research on pupils' perceptions of assessment for learning in religious education. Pupils can be seen to occupy various places on the spectrum from

task-focused to ego-focused in religious education, and this is coloured by more general educational concerns, but they can also be seen to tackle the task of reflection on wider goals of learning.

The chapter also contributes to the methodological rigour of the research by showing how the narrow research questions for the pilot study needed to be reconceptualised in order to deal effectively with pupils' experiences of learning. The design weakness in the pilot study led to a refinement in the main research. However, the next chapter analyses the data on generic assessment issues from the main research, as that follows on more directly from the data in the pilot study.

Chapter Eight

Self-assessment and pupil identity: analysing the main research

Having analysed the data from the pilot study in the previous chapter, the data from the main research can now be considered. The data is used to shed light on the *research* questions, and then these responses can be used to illuminate the *general* questions. The *general* question from the assessment literature was to examine classroom factors that affected self-assessment (chapter two). In this chapter the data that relates to this research question will be examined, and the next chapter will consider the second research question.

The first research question was: what classroom factors affect self-assessment? This was transposed into a research question, and was broken down into sub-questions (see section 5.2):

How can we understand self-assessment in the context of pupils' own ideas about their own learning and the classroom setting?

a. How do pupils' views on different self-assessment techniques vary?
b. Do pupils become more motivated as their ability to self-assess develops?
c. What else affects pupils' identities as learners?

The analysis in this chapter will focus on self-assessment in the context of pupils' own ideas about their learning in the classroom. This will be in four sections: firstly, a short analysis of pupils' understanding of self-assessment at the start of the research – did the pupils have a developed sense of their ability to assess themselves? This leads on to a discussion of pupils' attitudes to more summative processes, especially the use of levels. Thirdly, by way of contrast, pupils' views on the formative use of self-assessment are analysed. Lastly, pupils' attitudes to the formative use of summative processes are considered.

8.1 Pupils' understanding of self-assessment

There was a significant contrast between the pilot study and the main research in pupils' understanding at the start of the research of what self-assessment meant. Many pupils in the pilot study were able to explain self-assessment clearly (see above, section 7.4), but in the main research they were less clear. Only one could describe it:

> NF: what do you think I mean by self-assessment?
>
> Alan: Assessing yourself
>
> NF: In what way? How would you describe that process?
>
> Alan: Err. Dunno [C1]

A few understood it as the process of self-marking:

> Jake: um well you read through your work, and then you um, then you think if it's good or bad, and how good it is and you give it a level. [B1]

Only one pupil explained it as a metacognitive process:

> Ahmed: discovering what you've learnt, by commenting yourself, yeah commenting yourself
>
> NF: right, right, so it's comments...
>
> Ahmed: by yourself [D1]

The word 'discovering' is particularly striking here. It points to Ahmed's awareness of what he has learnt, as an act of metacognition, especially as he highlights it being 'comments', i.e. feedback, 'by yourself'. His was the most fully 'task-focused' view from the class at the start of the research.

This sets the scene for an analysis of self-assessment over the year. Unlike the pilot study, this class seemed to have little experience or understanding of it. They may have carried it out in lessons, but they could not articulate the process as the pilot class could. From a research perspective, it raises the issues of sample size; from the teaching perspective, it shows how different classes have their own dynamic.

8.2 Pupils views' of summative processes

It is therefore not surprising that much of the discussion during interviews focused on the nature and effect of the use of levels and marks from the teacher; pupils understood the term 'assessment' to be teacher assessment. A discussion about any aspect of assessment was bound to bring up a wider set of issues, as the pilot study showed (section 7.6.2). This was the model in their

heads. Moreover, it should also be borne in mind that these pupils sat major national tests (Key Stage 3 'SATS') towards the end of the year. These were in maths, English and science; the results would affect their setting in some subjects, notably maths, for the next two years of their schooling, for major qualifications (GCSEs) at the end of Year 11 (aged 16). Their explicit concerns with these were raised occasionally in the interviews. This example is from an interview when their SATS tests were about six weeks away:

> Alan: what's the test then in RE for the SATS?
>
> NF: there isn't one, SATS is only in maths, English and science
>
> Alan: I am sure someone told me there was an RE one. [C2]

These pupils were in the midst of the build-up to a critical summative assessment and were anticipating the start of a curriculum focussed on qualifications in the following year. Their particular assessment careers were likely to be shaped by these looming events.

In what follows, the data is divided into ego-focused responses and task-focused responses, as described in section 2.2 above (e.g. Gipps 1994, Dweck 1999). One might imagine that ego-focused pupils expect the teacher to apply levels and are unable to perform it themselves, because they see assessment as an external judgement about their innate ability. By contrast, task-focused pupils would be willing and possibly more able to self-assess their work, but would be reluctant to use levels because they do not give feedback. However, as in the pilot study, it was not as clear-cut as that, and other factors were also at work.

8.2.1 Summative processes: ego-focused responses

In many cases, this research echoes the widely established literature that shows that ego-focused pupils are dependent on levels or marks (e.g. Butler 1997, 1998, Gipps 1994, Black and Wiliam 1998a). Levels were seen by over half the pupils interviewed as essential for knowing their 'place', or as a target.

> Jake: Levels are cool
>
> NF: Why is that? Why are they cool?
>
> Mike: up the levels!
>
> *Laughter*
>
> Jake: um well, without them we don't really know, well I don't – I haven't a clue and so levels, levels are – a way of knowing how I'm doing. [B1]

However, the data also confirmed that using summative techniques, such as marks, grades or levels is de-motivating for those who do not succeed because they attribute their attainment to ability (e.g. Gipps 1994, Harlen and Deakin Crick 2002). Many had a fixed view of their capabilities, and would not believe that they could improve.

> NF: does having the level motivate you?
>
> Jenni: what does that mean?
>
> NF: does it make you want to do better?
>
> Jenni: no, because you just look at the level and think "oh well that's it". [C1]

It also confirmed the view that this was because pupils could treat the levels normatively, and compare themselves with the rest of the class (Gipps 1994, 41). They would judge their performance in relation to how well they had performed against their classmates, and use this comparison to judge their ability.

> Mike: it can, it can bring you down though because if you got a low level and your friends have got higher levels, >Ellie: yeah< then it makes you feel, um, stupid >Ellie: as if you're dumb< compared to them
>
> Lucy: and it makes you, like, um, give up
>
> Mike: it makes you see that there's no point in it >Lucy: yeah< because >NF: OK< because you're not going to get anywhere with it. [B1]

'Feeling dumb' was therefore often understood comparatively, and was demotivating.

Furthermore, the pupils also freely acknowledged that they could be motivated by comparative success – getting a higher level than their classmates. These pupils were still ego-focused, but were simply successful in the 'competition', and this could spur them to success:

> Billie: you can sit there and look really intellectual if you've got a high level >*laughter*< beard scratching. [D1]

> Alan: I like levels because when I get a high one I can gloat at Billie and say I got that level. [C1]

For Billie, a high level entitled him to show off within the class; he could act like a person with those qualities. Alan wanted to compete with Billie. Neither suggests that education could be collaborative, or co-operative.

Indeed, as research predicts, more successful pupils thought that getting a low level could be motivating:

> Kevin: I think if you gave me a level 2, it would make me want to do better, because level 2 is pretty poor
>
> Billie: there's a long way to go if you've got level 2, and you want something like a level 7
>
> Kevin: you'd aim for 6 or 7
>
> Billie: 6 or 7 yeah, yeah
>
> Ahmed: it makes you more determined. [D1]

Both Billie and Ahmed were likely to achieve these higher levels; given that the average expectation for the end of Year 9 was level 5/6 (QCA 2000), they were discussing comparatively high target levels. This matched Butler's finding that levels were motivating for the more able (Butler 1987, 1988) though for the 'wrong' reasons. Their generally successful assessment careers allowed them to discuss possible 'failure' in this way.

However, many pupils recognised that their experiences would fluctuate in this situation. The full context of Alan's 'competition' with Billie is revealing:

> Alan: I like levels because when I get a high one I can gloat at Billie and say I got that level >Sara: *laughs*<
>
> NF: so it's competitive for you, against someone else?
>
> Alan: yes. But if I lose, I don't like that. [C1]

For one pupil, even the sense of success was itself unpleasant:

> Aaron: I don't like getting my levels, personally...I'm not just saying that but you know, I don't like levels, cause I just don't like getting them, if they um, if they're too high I think "oh yeah, I'm really good, I'm a boffin!" kind of thing, and if I do too low I think "oh, I can do that better" and, and oh I just don't like levels, completely. I just don't, any subject. [C1]

Aaron here recognised the whole process as one of fluctuation, and preferred to have no part in it; as can be seen, he was very emphatic about this. Strikingly, he also sees this as being in 'any subject'; this is part of his 'assessment career' throughout his schooling. Aaron felt himself caught up in an ego-focused process, from which he wished to escape. He could not stop the feelings of elation or despondency, but he could recognise them as an unpleasant pattern. The terms 'too high and 'too low' also reflect this. He does not suggest that this is about social conformity; for instance there is no suggestion that he will be teased for being a 'boffin'. Rather, it is in relation to his perceived sense of his own ability that they are excessive.

8.2.2 Summative processes: task-focused responses

By contrast, however, a few pupils recognized that there could be formative aspects to the use of levels by teachers. These pupils valued levels within the context of a positive view of their potential from their teacher:

> Dave: with um if, you got say a level five, and the teacher comes along and says "oh, I think you could get a level six, you should be working towards that" and you'd be like "oh, he thinks I can do it" so that's an extra, like, boost for you to do better [A3]

Dave here understands that a 'social constructivist' teacher will enable him to succeed in his 'zone of proximal development', rather than one who regards his level as a fixed measure of intelligence (Vygotsky 1987, 210). This illuminates the suggestion that pupils' perceptions of their teacher's learner-centred practice are critical (Deakin Crick *et al.* 2007, 276, Marquez-Zenkov *et al.* 2007).

Moreover, one group of interviewees distinguished between its different roles in tests as opposed to homework, and separated the teacher's different roles:

> Billie: I think on homework comments are better than levels
> Ahmed: yeah, I think on tests levels are better than on homework
> Billie: if you're just going to put one level that shows the teacher is not that bothered, really...
> Kevin: if, you've got comments, then job done. [D1]

This exchange does not of course suggest that pupils automatically viewed levels formatively, but is shows that they are able to distinguish between the different functions of different forms of assessment, and indeed that the teachers have a variety of roles. They recognise that there is a summative role, in which levels are appropriate, and a formative role in which they are not.

8.2.3 Summative processes: conclusion

On the basis of this evidence, it is too simplistic to say that the use of levels by teachers is *entirely* negative; however for the majority, the negative effects often dominated. There are four main factors. Firstly, it is affected by the inherent features of the task; clearly, some assessment processes are more summative than others, such as the use of levels. Nevertheless, secondly, pupils' assumptions about the role of the teacher or the pupil as assessor can affect this, and if pupils understand the teacher's approach as pupil-centred then they could understand the levels formatively. Thirdly, the behaviour of the rest of the class mattered, e.g. the risk of judgement on one's character as big-headed, or having classmates calling you names. Lastly, it is affected by the pupils' beliefs about their ability to succeed, or whether intelligence is innate.

8.3 Pupils' views of self-assessment

If these factors influence pupils' attitudes to summative processes, do they also influence formative self-assessment processes? The research that distinguishes between ego-focused and task-focused pupils tends to suggest that pupils' attitudes switch from one to the other as a result of the inherent functions of the

assessment techniques; thus in Butler's research, the pupils who only received comments had become task-focused *because* of this. However, Ecclestone and Pryor suggest that pupils will continue to react to assessment processes according to their previous experience, and this will therefore colour their engagement in these processes (Ecclestone and Pryor 2003).

The following analysis is divided into four sections. To start with, the pupils' attitudes to the nature of their own ability are discussed. Then their attitudes to the role of the teacher are considered, particularly the inter-relationship between teacher feedback and self-assessment, which leads to an analysis of how pupils grasp the criteria for self-assessment. Finally, the socio-cultural dynamics of the class are discussed.

In the following analysis, pupils commented on particular strategies that they had used in the lessons. This however is not to suggest that self-assessment is simply a collection of techniques, as identified in the comparison of constructivist and behaviourist approaches to assessment and learning (section 2.5), and in the pilot study. Rather, the process is embedded in and the product of these techniques, as the pilot study showed. The challenge for teachers is to integrate the techniques into their practice successfully, in order to engender the changes in pupils' attitudes to their responsibility for their learning (Black *et al.* 2003, 28, Laveault 2007).

8.3.1 Task or ego-focus in self-assessment

One of the problems with introducing formative techniques is that some pupils may not value them, especially if they are used to more summative assessment processes. They would therefore not engage with these processes, or only engaged in a superficial way. This can be seen in their comments on two techniques: the use of 'traffic-lights' and plenary lesson reviews (see section 5.4.1). Over three-quarters of the pupils immediately compared this to the use of levels:

> Ellie: well, I think that if we get told our levels, it's better because then we know what we need to improve...and so, like when we get our SATS levels, we know what we're aiming for, which is usually the one above, yeah [B1]

Ellie suggests that she needs to have the level as a target because without it she could not unpack the criteria, even though these techniques would give her the criteria without the level. Under a quarter of pupils related their use to revision for tests or exams;

> Helen:...I suppose if you're ticking it you could go back for your GCSEs, and you could actually go "oh yeah" [A1]

> Mike: I think it's an alright idea, because if you're revising, you can see the ones with the cross which are the ones you should work on first, and then go back to the ticks, so if you went through your book, and found all

the ones with the crosses you could revise them first and then do the ticks. [B1]

These show how pupils appropriated and judged the effectiveness of the techniques according to their assessment careers. They were used to levels and revision for tests; but although I used levels once a term, I never gave them tests, so Mike was considering their worth within his wider experience of assessment.

Nevertheless, other pupils could clearly identify formative benefits to levels, and the way that it encouraged a sense of progress in their learning. On reviewing the lesson:

> David: ...I think it's good because you know whether you've understood what's gone on in the lesson, so...you've learnt, you can say that you have at the end of the lesson...the main thing is you can walk out of the classroom in the morning and say to yourself, I've actually achieved something this lesson. [A1]

> Alan: that's good because then, you know, like you can put it at the start and at the end then you know >Aaron: umm< like where you've improved
> Aaron: you see, I agree with that >Alan: what you need to improve on< because it's easier because you know what you need to improve on [C1]

For David, the process helped raise his self-esteem; he could appreciate that he had made progress that lesson. This highlights the meta-cognitive dimension to self-assessment, similar to Ahmed's remarks about discovering what you know

(section 8.1). For Alan and Aaron, the process helped them identify both progress and areas for improvement; it made them more task-focussed.

However, some pupils felt that traffic-lighting was not formative, similar to pupils in the pilot study.

> NF:...does it help you know where you are?
>
> Lucy: it helps me know where I am, but the module traffic-light sheet ...doesn't help me learn...'cause <u>learning</u> is like getting more information, that's learning, like in projects and that sort of stuff
>
> NF: and it's not helpful when you do [*the traffic-light sheet*], directly?
>
> Lucy: It helps me know where I <u>am</u>.

Lucy did not feel that it gave her clear enough guidance on what to do next. She echoes Sadler's insistence that to be genuine self-assessment, the process must engage in an action that leads to closure of the gap between current attainment and the desired goal (Sadler 1989, 121).

Finally, however there were some pupils who simply carried out the activities because they were told to by me, their teacher. As in the pilot study, theirs was procedural compliance:

> Jenni: ...it's just a waste of, well technically, it's just a waste of time. It's just, you don't really care what you've done in the lesson. It's just like the end of the lesson, the lesson is over so yeah >Sam: you want to go< yeah, you want to go, so you don't care what you've done, so. [C2]

Throughout these examples one can see different pupils interpreting the techniques differently. For some they represent a threat, for others they are a boost to self-esteem; for some they are meaningless, for others they are a powerful way of appropriating their learning; for some they are worse than the use of levels; for others they are better. Not all classes will have this particular combination, but it is reasonable to suppose, especially in the light of the pilot study, that these elements in some combination are common to most classrooms. Thus the practical problem for any classroom teacher is likely to be that different pupils will respond to different self-assessment techniques in different ways depending on various factors including the nature of their assessment career: are they task-focused or ego-focused? Are they concerned with popularity? Do they enjoy or are they embarrassed by getting higher levels than their peers?

Moreover, pupils may not progress towards a more formative approach to assessment in a straightforward way. The new tasks themselves just become part of the pattern of school life, and it is not always easy to maintain a serious attitude towards learning:

> Penny: but it is always like at the beginning, I kind of think about it, but at the end I just go "done" and I just kind of mark in green, I just go green, green, green, green, green. *laughs*. Even if I don't, I don't actually read them. I just
>
> NF: it is just like a process that the teacher is asking you to do? >Poppy: YEAH< so it is just like

Penny: I don't really think about it [B3]

This highlights both the fluid nature of pupils' learning identities in the classroom, as a mix of being task-focused and ego-focused, as well as the nature of compliance in school.

8.3.2 The teacher and the pupils

The teacher's ability to manage this transition in pupils' attitudes successfully is also dependent on pupils' perceptions of the teacher-pupil relationship. If they continue to think of the teacher's assessment role as an essentially testing one, then they are unlikely to engage in formative processes. It is not just that pupils must come to a new understanding of their own learning, they must also come to a new understanding of the role of the teacher, and the nature of classroom activity. Perrenoud describes it as 'reconstructing the teaching contacts' (Perrenoud 1991, 92 in Black and Wiliam 1998a, 21).

This is particularly important in the light of Vygotsky's remarks about the importance of 'collaboration' between pupil and teacher in order to aid progress (Vygotsky 1987, 210). If pupils do not feel the relationship to be a collaborative one, then they are unlikely to respond collaboratively (see Deakin Crick *et al.* 2007). The practical problem is if the pupil persists in seeing assessment as being of himself or herself, rather than of the work.

Some pupils were able to discern the formative teaching relationship, and identified particular responsibilities for the teacher and pupils in this.

> Alan: and if you, if you self-assessed yourself too much then um - there'd be no point in having a teacher. [A1]

> Ellie: well, I think that you, you should sort of like explain on the board and kind of like write it out and what we're got to do and then we'll check it. That's what we do in most of our other lessons. [B2]

Ironically, Ellie seems to be describing precisely the lesson review process that I believed that I had implemented; however it may be that she wanted more detail in terms of the *tasks* in the lesson. Others struggled to explain the reasons for the techniques:

> Chas: ...so you put down, say, that's red that you don't understand and green and stuff like that? Yeah, I think that's because – 'cause you look after this, and you look at this afterwards, don't you? And then instead of us telling you, you've got it down on a bit of paper what we understand and what we don't understand sort of thing and you can just keep it. [A1]

In this example, Chas tries to contextualise the use of the traffic-light sheet, but he does so in a way which passes the responsibility back to the teacher. Self-assessment for him seems to serve a diagnostic function for the teacher, rather than being of benefit for him.

One area on which the pupils gave me clear advice was the failure of the 'rapporteur' technique (see section 5.2). As explained, this was because pupils treated the announcement of the plenary as a signal to pack away, i.e. as the end of the lesson. I found that however much time I tried to allow, pupils still responded the same way, so I abandoned it in November. Later, in July, pupils reflected on this:

> Penny:...The thing is like people get – really like annoyed if you start doing something new at the end of the lesson, like, if ever it is like – the end, if it is like 9:35 then everyone's is like >David: yeah< "oh yeah time to get, start packing up", and you're like "Right, sit down, now we are going to do this, and do this" >David: yeah< and then everyone gets really annoyed because, and no-one listens they just start packing up and like standing up behind their chairs. [A3]

However, they also suggested solutions to the problem.

> David: I think with you doing it, you need to reinforce it more as part of the lesson because at the end of the lesson, everyone was packing up. Everyone was paying attention on getting outdoors and don't care less, and it's a very harsh way of putting it, but it's... if you were to bring it forward in the lesson – um, just before everyone packed up, then I think it would be more – accomplished. [A3]

> Kevin: if we had started it in Year 7, then we would have found it much easier, but the fact that it was started in Year 9, then it was a bit of a culture shock. [D3]

David's comment is illuminating because he could see its value and therefore he recognised the need for the teacher to build the process into the lesson. Kevin's suggestion is less of a solution to the immediate problem, but does show that pupils recognise how building formative processes into their assessment careers could be a lengthy process (Learning how to Learn 2002). They were unused to this kind of activity, and expected to simply stop what they had been doing and unreflectively leave the class. It also shows that techniques are not formative simply because the teacher intends them to be, as Perrenoud had once suggested (Perrenoud 1988). Incidentally, Kevin's remarks support the view of the Learning how to Learn project that this process is best done when pupils are new to a school (Learning how to Learn 2002).

8.3.3 Understanding the criteria

If pupils need to develop a new understanding of the role of the teacher, they also need to develop a better understanding of what it is they are to do. The tension in the literature on self-assessment between those who see it as a process of enculturation into pre-existing criteria (e.g. Sadler 1989, Wiliam 1998), and those who consider that it should be more autonomous (e.g. Pillonel and Rouiller 2001) has been highlighted. Part of the research focused on the first process, particularly through the use of the lesson review, the module review and the use of the level-path sheet.

My module review sheets included both details of progress in learning about religion ('I can describe *Yad Vashem*') as well as learning from religion (e.g. I can relate the Jewish need to remember with my own experiences', see appendix 1). What did the pupils make of these techniques?

Some discussion in the interview focussed on how specific the wording of these sheets was, and in particular the pupils compared my sheets with those they had used in science. Some students preferred the format in science.

> NF: and are they useful in science? =
>
> =Alan: yup >Aaron: yeah< they are more useful in science.
>
> Sara: um hum...
>
> Aaron: there's no kind of correct thing to base it on really. There's your idea, and your idea only, kind of thing...In maths and science you've got a definite answer, like 1+1=2, whereas in RE, it's like 'do you know more about?' > Sara: it's basically your opinions< Do you believe?' [C3]

Others however preferred my format. This was firstly because they were shorter, which helped them focus on them.

> David: yeah but with the science one, if you see, like forty-odd questions, you are like 'oh I can't be bothered to go through them' Green, green, green, green, green, green, green. With yours > Penny: yeah but < you think its short and oh, you know, 'I might as well look at it'
>
> Dave: you get to think, you read it and you think about it. [A3]

Secondly, they liked the open questions, particularly in terms of learning from religion and more personal outcomes.

> David: you could do with an opinion box next to all your questions
>
> Penny: you can't say, are you more tolerant towards Muslims or are you not? Because it's quite a long answer. [A3]

In these discussions, a number of issues concerning sharing criteria with the pupils can be seen. There is a practical question about how specific the wording of any self-assessment review sheets should be, but this is inextricably linked with a wider discussion about the nature of the subject itself. As we have seen, one of the learning outcomes in religious education should be the ability to evaluate and reflect on the content for oneself (learning from religion), but this skill is transferable, and hard to define. The criteria are 'fuzzy' and they are 'not reducible to a formula that can be applied by a non-expert' (Sadler 1989, 124). A range of answers could be 'good', and the quality required is harder to identify in a more general way.

Moreover, Laveault points out that pupils fail to grasp learning objectives not only when the criteria are not specific enough, but also when they are too rigid and therefore too constraining in the long term (Laveault 2007, 216-217). My view as a teacher was that pupils needed to grasp the *general* criteria against which their work was to be compared. Rather than simply monitoring their progress in one topic, they needed the framework to make sense of their progress across all the topics; what were the generic skills that they needed to grasp? Secondly, if one of the attainment targets was the ability to learn from religion, then pupils needed

to self-assess that. There is a balance to be struck, or rather perhaps this sharing of criteria also needs to be differentiated.

8.3.4 Socio-cultural dynamics

Pupils' experiences of assessment are also inextricably entwined with the socio-cultural dynamics of the class. Summative processes that tend towards norm-referencing will have an immediate impact because of the inevitable 'social comparison' that results (Gipps 1994, 41). The conclusion is that this is demotivating for the less able, but is motivating, albeit for the wrong reasons, for the more able (e.g. Gipps 1994, Black and Wiliam 1998a, Harlen and Deakin-Crick 2002). The data on peer assessment particularly illuminates this issue. This involved looking at examples of other pupils' work in two ways, as described in section 5.4.1. One of the peer assessment tasks was to use the level criteria on three examples of work from pupils in the class. Clearly the earlier discussion on the use of levels for self-assessment is relevant here (section 6.2.3). The other was to comment on examples from another class before attempting their own work.

About a third of the pupils expressly viewed this in a task-focused way, but a quarter treated the levels as a competition. Peer assessment using levels was vulnerable to the norm-referencing effects of levels. Thus:

> Jake: yeah but you're not going to get, you're not going to give someone a - like a 3, even if it is quite - not that good because um - it's just not, you wouldn't want to get unpopular with that person. [B1]

Jake would not publicly award a low level to another pupil in the class because of the impact on social relationships. In terms of other formative peer assessment techniques, notably making comments on other pupils' work, two broad issues emerged: whether peer assessment was helpful or not; how it impacted on the social dynamics of the class.

Three quarters of the students, representing a broad range of general attainment, thought that peer assessment was a helpful process.

> NF: ...What about looking at other people's work to do that? - Is it helpful?
>
> Ahmed: yeah, I suppose a bit, because you get lots=
>
> =Annie: yeah I think so...see how they've taken it (**)
>
> Kevin: it's interesting [D2]

> Chas:...It gives me some ideas, if you know what I mean. I actually understand it more. [A2]

However, what concerned a couple was the issue of originality; if they looked at another pupil's work before they attempted the task, their own work might be a copy. The process of judging originality is not easy; it is a highly qualitative 'fuzzy' criterion (Sadler 1989, 124), which means that it is not reducible to a formula, and it is not clear what is acceptable or not. A few were very scrupulous:

Lucy: it did help but then like, it was quite hard not to put what they had put, if you know that I mean, that's what I am saying, if you understand, isn't it.

NF: and is that bad?

Lucy: well yeah, if it's not your own work [B2]

The process of comparing the quality of one's own work with another pupil's without compromising originality was tentatively expressed by one pupil:

Dave:...it would [help] if you gave Kate's script to someone who had got a A, got an A, writing in that style...it's just, if you read all her pieces, all their pieces out, as I said, if she allowed, er, it would give you an idea of what you should be writing, or could be writing...and that would give you a thing to hold it against, but it would still be original. (*Like when you used Kate's) to see what she has to say by what you have already written in that sort of...format. [D2]

The expression 'a thing to hold it against' captures the benefits of peer assessment. Dave understands that the other pupil's work is not a model to be replicated, but it acts as an exemplar of the criteria. However, the helpfulness of peer assessment could depend on its standard:

Ahmed: it only helps if the writing's as good as mine

Annie: umm

Ahmed: if it's a bad piece, then it does not help, it just makes for laughs

Annie: if it's bad, you could help the other person

Ahmed: but if it's good then you get more, better. [D2]

Ahmed seems to combine a task-focused ability to make use of peer assessment of what he considers to be suitable quality, and an ego-focused ability to compare his own abilities with others.

This example highlights the other aspect of peer assessment: the social dynamics of the class. A major factor in this was the potential embarrassment if one's own work was the subject of the peer assessment process.

NF: right. So the embarrassment point is quite a big point?

Lucy: yeah >Mike: mmm<

Ellie: I think they were all embarrassed when=

=Jake: I wouldn't be that embarrassed

Ellie:...I would. [B1]

NF: Is it useful to look at...other people's, not necessarily in the class, other people's work?

Annie: yeah

Billie: yes, just don't choose mine. [D1]

Billie's remark highlights the tension clearly; he could see the educational benefit, but he was also fully aware of the social implications of this. Several pupils were happy for their work to be chosen, for instance Jake in the first example above. In the next example, Dave's work was one of three pupils whose work had been handed round for pupils to assess before assessing their own:

NF: OK. Um the last thing was using, when ...I gave you the examples of other people's work, so we did what's >Helen: oh yeah< so we did what's called technically called

Helen: oh yeah, it's supposed to be anonymous, but we went "oh Dave, look! Al, look! It was Kate's!"

NF: was that, was that a helpful thing to do?

Dave: I think that was quite good.

David: yup. [A1]

Many pupils thought that the solution to this was to use anonymous work from a different class.

Ahmed; or, er, you could get a piece of work from the past >Kevin: right, yeah< like a year or two before. [D2]

Mike: It should be, like, work from other classes > Ellie: yeah< so you can't tell. Because, if it's in the class, you can tell and it would be embarrassing for the other, for the person. [B1]

Even so, a few had concerns even with this.

Sara: That was better than having someone's from your own class, but they still might not like that=

=Alan: I still think it's an invasion of privacy, because its still someone else's views and it should be for them and the teacher, not other people to look at, 'cause they wrote it

Aaron: they might not want it to be=

=Alan: they might write something personal. [B2]

This reveals their concerns about the nature of the teacher-pupil relationship. Written work is perceived as a private matter between the pupil and the teacher, not something that can automatically be shared with other pupils. Thus even if pupils were to understand peer assessment in a constructive way, it may still be seen as a breach of trust, and therefore undermine their perceptions of the teachers as genuinely pupil-centred (Deakin Crick *et al.* 2007, 303, Marquez-Zenkov *et al.* 2007, see also Fancourt 2007, 63).

8.3.5 Pupils' views of self-assessment: conclusion

In this section we have examined four different factors that affect the pupils' ability to make the most of the opportunities for self-assessment. The first was the extent to which they treat self-assessment in a task-focused or ego-focused way; the second was their understanding of the role of the teacher; the third was their views on the ways that they could make sense of the assessment criteria; and lastly the links between peer assessment and the social dynamics of the class.

As in the pilot study, Laveault's general principles of self-assessment were illuminated (see section 7.6.1, Laveault 2007). What is striking for a teacher wishing to develop self-assessment is how a number of different factors interlock in a complicated way. Thus, the use of peer assessment could be either unhelpful because it creates embarrassment and stifles originality, or helpful

because it gives pupils 'a thing to hold it against'. The use of general descriptors of criteria could be helpful for some but unhelpful for others. It may also be difficult to implement some techniques, such as the use of the rapporteur, because of both embarrassment, and because pupils do not respond to it effectively. Overall, this shows that effective self-assessment is embedded in a range of formative practices and cannot be developed without them, but the different elements - feedback, sharing criteria, and the use of levels - have different functional characteristics in different situations with different pupils.

8.4 Pupils' views of self-assessment using levels

This research has shown how four main factors affect pupils' attitudes to summative teacher assessment and formative self-assessment: their understanding of the nature of the assessment technique, their own self belief, their perceptions of the role of the teacher, and the socio-cultural dynamics of the class. So far, the research has borne out the broad distinction between, on the one hand, ego-focused attitudes and summative processes and, on the other, between task-focused responses and formative self-assessment. However, it also appears that this is not a clear-cut dichotomy in pupils' minds; instead, pupils' individual assessment careers will colour their attitudes to and therefore experiences of both summative and formative approaches.

Self-assessment is generally seen as being a formative practice, but a critical problem 'relates to the possible confusions and tensions, both for teacher and for learners, between the formative and summative purposes which their work might

have to serve' (Black and Wiliam 1998a, 59, see also Black *et al.* 2003, 30, DfES 2003, OFSTED 2003).

In my research, pupils were asked to apply levels to their own work, but how successful was this as a formative process, or did its potential for norm–referencing dominate? Strikingly, Maqsud and Pillai found that the experience of self-scoring led pupils to be less hostile to the use of marks (Maqsud and Pillai 1991). Thus even the transfer to the students of the basic process of self-scoring, with no element of formative feedback, reduced their hostility to summative teacher assessment. More specifically, was self-assessment using levels more like summative teacher assessment or more like other forms of self-assessment, or was it different again?

8.4.1 Self-assessment using levels: ego-focus

Some of the processes in the research involved self-assessment using levels: the use of a target-setting sheet to judge their own work against national levels (the 'level-path sheet', see appendix 3), and a peer assessment exercise using the level criteria. The first problem was simply carrying out the process.

> Lucy: it was quite hard because when they describe what you had to have for the level, I didn't know whether I had that. [B1]

> Jake: if it's, if it says something like you can understand, Paul's religious experience, and get ideas from it but – I don't know I wasn't sure about

that and er – I just went for level five because it was the one in the middle. [B1]

These responses highlight the issue of 'plasticity' (Wiliam 1998). It is hard to for pupils to get a sense of the kind of quality that a descriptor represents; it is harder to then apply it to their own work. Jake felt that he could do this if they were very specific, but could not if they were more general. For some this made it pointless:

> Helen: yeah, but that isn't official and that's what really (*matters)
>
> Jenni: that was just a waste of time
>
> Sara: waste of time, and a waste of ink...because you just have to take a random guess. [A1]

Sara feels very ill-equipped at this. Further, it is perhaps no surprise that some pupils approached the use of levels in self-assessment in a normative way; they were ego-centred (Dweck 1999).

Importantly, the simple fact of over-estimating was itself de-motivating.

> Jenni: it was like...if you give yourself a level 5 and like get a level 4...and then you get like a lower, you feel like really down and stuff. [C1]

Both these quotations show the emotional impact of assessment clearly. To ask pupils to practise self-assessment in this way is to ask them to judge themselves in the sort of way that they expect teachers to judge them. In Pillonel and

Rouiller's phrase, it is 'internalised hetero-assessment' (2001, 3). Therefore, if they find that the levels that teachers give them are de-motivating, then they will also find self-assessment to be de-motivating. The internalisation of their experience of the assessment process also runs the risk of being doubly uncomfortable because not only is the pupil's work 'actually' at a low level, but also the pupil has failed at the process of self-assessment, by getting the level 'wrong'.

On the other hand, to give oneself a comparatively high level was to label oneself clever, and this ran the risk of being ridiculed within the peer group.

> Mike: I'd say, I'd say a lot of, er a lot of people wouldn't say what they thought in case they thought: 'Oh I think I've got a level 6 but I'm not going to say it because people are going to shout stuff at me'. [B3]

Thus even for successful ego-focused pupils, the process was not a motivating one. Against the literature on ego-focused motivation, one could balance the literature on school sub-cultures, in particular on the negative labels that some pupils used to describe successful pupils: Willis's "ear-'oles" (e.g. Willis 1977). The label of 'intelligent' from high marks could become a self-negating prophecy as much as a self-fulfilling one. For many pupils, the social climate of the class did not allow them to stand out, whether as being 'dumb' or being a 'boffin'.

These concerns were also often linked with social concerns about how one would be perceived in the class if the self-awarded level was 'wrong' as against the teacher's 'real' levels:

Ahmed: you might do yourself a mark wrong, and that would not be helping at all

Annie: people might think you were big-headed

Kevin: if you give yourself a level that's above...

Ahmed: or if it's below

Kevin: you might look a bit of an idiot. [D2]

Annie, Kevin and Jake here highlight concerns about how the rest of the class would perceive them. Kevin in particular emphasises the risk of humiliation. A way to avoid disappointment and humiliation was to be harsh on oneself, but this of itself was also de-motivating:

Sara: because most people, well, I don't feel I do very well, so I set myself like really low kind of standards, so I really think I am doing really bad, so it doesn't really motivate me to try and do well in my work, because I think I am going to get so low. [C3]

This is a kind of self-protection; in order to avoid disappointment Sara has lowered her self-esteem.

One interviewee explicitly recognised that this was a question of the pupil's self-esteem:

Penny: I think people with low self-esteem will probably just say 'oh no I did rubbish', if they think that was rubbish, they are going to give themselves a low mark. [A3]

This was particularly an issue in relation to the use of levels for target-setting. Over half found this a fraught process:

> David: yeah, because say one day you are feeling pretty cocky, I mean what you do is you come in and you say 'right I am going to be level 6', you give it to the teacher, and the teacher goes 'right, I am going to expect a level six from them because they think they can do that'. Then you come in another day and you are really nervous and you just don't think you are going to do well at all, and if they've put pressure on you because you have to do the level 6. [A2]

One of the most striking features is that there is only one example of self-assessment being a *positive* influence on ego-focused pupils. This is because ego-focused pupils respond to marks that they receive, and therefore need the assurance of an external view. They rely on the external judgement of success. However, the exception is striking:

> Alan: that's why you always assume that if you put a three, and you get a seven you're clever. [C1]

Alan wants to manipulate the system so that he gets the ego-centred praise: being 'clever'. By consciously under-estimating his performance, he believes that he can artificially generate this praise, despite the self-deception that this would involve. Another pupil, Kevin, quoted above, had suggested that if you unintentionally underestimated your performance, others might consider you 'an idiot'. Alan does not consider this: it is not clear how successful his strategy

would be. What is however worth highlighting is that a pupil can be so adept at devising strategies for coping with an ego-focused system.

In general, therefore, almost all pupils felt that the teacher alone should apply levels:

> Lucy: yeah, but I prefer teacher
>
> Mike and Hannah: yeah
>
> NF: you like the teacher to give you the levels?
>
> Mike: the professional view
>
> NF: rather than doing it yourself
>
> Mike: yeah. [B3]

For these pupils, self-assessment using levels was an awkward process. They could estimate it too high, which was simply unhelpful because it was wrong, and would be de-motivating once it was corrected by the teacher; it would also be socially embarrassing if other pupils were aware of your error. Moreover, if you were to estimate it too low, that would be de-motivating itself, and 'wrong'; it would also be embarrassing if other pupils were aware of your error. It is unsurprising that many preferred the 'professional view', and it highlights the importance of other socio-cultural factors that will influence whether pupils can use levels formatively.

8.4.2 Self-assessment using levels: task-focus

At other points however, some pupils recognized the potential formative benefits of using levels in self-assessment, particularly for target-setting. Thus the use of levels for self-assessment could also be a formative process.

> Dave: when you mark yourself, on that sheet, it shows, it has all the things that you need to do to get that level, and you know if you've got that level then you know that you can express different ideas and do what it says to get that level. [A1]

For this to be successful, about a third of the pupils recognised, or assumed, that they would know what the level represented. Levels could be formative if the pupils understood that the criteria were shared. Some pupils considered one particular technique helpful in this respect – the 'Level path' sheet (see appendix 3).

> NF: is it useful having the levels on it? Is that helpful or not helpful?
> Dave: it gives you a feeling of self-achievement. >Helen: yes< so that you can say that "I'm doing good but I can do better"
> David: you can use it when you sit down to write your piece of work
> Dave: "I'm doing bad and I can do a lot more".
> David: and you can look at this sheet and say right to get a level 3 I've got to do this, or to get a level 7 I've got to do this, so I'm going to go, what shall I do? I'll go for level 7 and try and do this, but at least if I aim for a

level 7, I won't be disappointed if I get lower >NF: right< because I will still have done my best. [A1]

If they knew their existing level, then a sheet that helped them identify what they should do next was helpful; they say that they can 'aim', 'improve' 'try for this'. At the end of the last example, David expresses this eloquently when he recognises that 'a feeling of self-achievement' is not linked to the level but to doing one's best; this is task-focused thinking.

8.5 Answering the research question on assessment processes

In reviewing this section, we can return to the research questions on assessment:

> How can we understand self-assessment in the context of pupils' own ideas about their own learning and the classroom setting?
>
> > a) How do pupils' views on different self-assessment techniques vary?
> >
> > b) Do pupils become more motivated as their ability to self-assess develops?
> >
> > c) What else affects pupils' identities as learners?

We can see how pupils' views on self-assessment seem to vary in accordance with their identities as learners, and in particular with their assessment careers. A range of attitudes, from task-focused to ego-focused, are represented, but with

most being a messy mix of both; in particular this affects, and is affected by, their views of the individual techniques, notably the use of levels. These assessment careers extend across all areas of the curriculum. However, it is no surprise that for most of these students most of the time, summative classroom assessment was a negative experience, as other research shows (e.g. Butler 1987, 1988, Black and Wiliam 1998a, Harlen and Deakin Crick 2002). This was possibly exacerbated by the fact that they were in one of the most significant school years for summative assessment, in a 'high stakes' educational environment (Looney and Wiliam 2005). They were haunted by concerns about levels throughout the research.

Nevertheless this is not to argue for a kind of assessment determinism. All the pupils in the research were able to explain their approaches to assessment, and make judgements about their preferences and choices; for example, many compared their experiences of traffic-light sheets in science and in religious education. They could look back on the year in the interviews. As Ecclestone and Pryor suggest, this is a habitus that the pupils both create and which creates their understanding of their role in school. It is hard from this research to know to what extent my efforts in one subject affected their perceptions of assessment elsewhere. Did the use of traffic-lights for module reviews help them in their understanding, and use, of them in science? Did all this soften the negative impact of summative processes in lessons generally, and more specifically the pressure of SATS tests?

If we now turn to the issue of motivation, the question is much harder to answer. This is because it is far from clear that their capacity to self-assess had

developed. After all, it would be extremely hard to measure. Further, it is far from clear whether their motivation had improved; again it is not evident from this data. It would be hard to be confident about the nature of the causal links between the two; even supposing that their motivation had improved and their ability to self-assess had developed, one would have to be clear that the latter had caused the former, and that there was neither another factor nor that the former had caused the latter. The pupils clearly commented on changes to their motivation as a result of the different techniques (see sections 6.2.1 and 6.2.2 above), but it is difficult to know fully how these interrelated over the year in religious education, and across their assessment career.

In answer to sub-question (c), on other factors, one key element is the learning dynamic of the class as a group. For instance, my decision to abandon the use of a rapporteur was based on how, as a teacher, I saw the majority of the class responding. For some pupils this may have been a useful technique, but they were not to experience it because of my approach to other aspects of the regulation of learning. This indicates the second factor, which is the pupils' perceptions of and relations with the teacher: do they see the teacher as the external judge of ability or as a facilitator of their learning? This may well be linked to more general feelings of trust in the teacher (Deakin Crick *et al.* 2007, Fancourt 2007). A third related factor is the social dynamic of the group – who is friendly with whom? Who is disliked? Are pupils sensitive about their work shared? The practical teaching implications of this were similar to the pilot study; self-assessment itself should be differentiated, and should become the object of a progressive training, as Laveault suggests (Laveault 2007, 232-233).

8.6 Conclusion

This chapter continues the discussion from the previous chapter on the way that pupil identity is shaped by and shapes classroom assessment, which is an important and significant issue. The data on pupils' perspectives has been analysed in the light of the research questions. First, their lack of a clear sense of self-assessment was identified, especially in comparison with the class in the pilot study. Next, their attitudes to the use of levels by the teacher were reviewed, which generally confirmed the widespread research evidence that many pupils found this a fraught process as it encouraged ego-focused attitudes to learning. However, there were a handful of examples of students who could approach levels in a formative way, or who could distinguish between different roles. Thirdly, the data on purely formative processes was reviewed, which showed that intended benefits of this were often mitigated by their ego-focused responses; thus it is too simplistic to judge the action functions of techniques simply by the nature of the technique or by the teacher's intentions alone. Lastly, the data on the formative use of summative assessment, i.e. getting pupils to use the levels themselves, was reviewed, confirming the mixed nature of responses to this process as well; pupils alternated between ego-focused and task-focused responses. This analysis has then helped to address the issues raised in the research question, and sub-questions.

Overall, it is original in shedding light on the general question of the classroom factors that affect self-assessment, by showing 'the messy point of teachers and pupils responding to each other culturally in relation to....assessment practices' (MacFadden and Munns 2002, 364). It shows how ego-focused and task-

Chapter Nine

Reflection in religious education: analysing the main research

In this chapter, the pupils' views on both the educational goals of religious education and their own progress are analysed. The general question on the pedagogy and assessment of religious education was whether the current attainment targets were aligned with all the learning outcomes satisfactorily (see section 3.5). However, the research question was re-formulated more specifically in terms of the pupils' views of learning outcomes and progress (see section 5.2):

1. How do pupils consider that self-assessment helps their progress in religious education?

   a) What do pupils consider to be progress in religious education?
   b) Do pupils think that they have improved at this?

The answers to these research questions will inform the more general question because in deciding what appropriate attainment targets and learning outcomes are, it is vital to consider what pupils themselves think is possible.

The pupils' answers to these questions will inevitably be conditioned by their experiences of religious education both with me and previous teachers. My approach was implicitly interpretive and dialogical (see section 5.4.1); it is therefore possible that the pupils' responses will mirror this. Nevertheless, the pupils may also bring other experiences to bear when reflecting on these questions, from other schools, and from outside school; furthermore it is

important to know whether they did in fact agree with these pedagogical ideals and assumptions – they might reject them.

Firstly, the views of pupils who rejected any value for the subject are considered. Then pupils' comments on 'learning about religion' are reviewed before turning to the range of views on different aspects of 'learning from religion'. The data shed light on debates about the usefulness of this distinction in the subject (see chapter 3, and e.g. Wright 2004, Jackson 2004a, Copley 2005, OFSTED 2007). Lastly, their views on the development of attitudes and values are discussed, and the implications of this for the formulation of policy documents.

In addressing these research questions, however, rather than dealing with each sub-question individually, the analysis is divided up in relation to the different learning outcomes. This is because at the coding stage it was often hard to separate the question of what the learning goal was from whether pupils had progressed towards it; to say one had improved at a particular aspect of learning was almost inevitably to recognise that aspect as a learning goal. Thus, any data relevant to sub-question (b) was also relevant to sub-question (a), but not *vice versa*. Examples of how they articulated each learning outcome are presented and then how they understood their progress towards this.

9.1. No value in religious education

The first important point is that at times a few pupils during the research process expressed the view that religious education did not have any value or they did not understand its value; they did not think that it had worthwhile or

relevant learning goals. One, Annie, associated the purpose and value of religious education with religious convictions "I don't believe in religion, I don't really care" [D3]. One pupil was rather more expressive at the start of his first interview:

> Alan: not really, um I don't see the point in RE... because it's totally crap. I mean, why do I really need to know what some guy on the other side of the world is doing, and why he is doing it? [C1]

For Alan though, it was not his lack of belief that made it irrelevant, but the remoteness of the lives of others to his own.

A couple were prepared to do what they were asked to in lessons, but did not have a wider understanding of what the subject was about.

> Sara: I don't actually have a clue what I need to do, to get better, but what I am doing.
>
> NF: right, so you kind of do what you are asked to do and
>
> Sara: that's about it really. [C3]

Sara is an example of compliance, dutifully doing what she is told but without any sense of it being *her* education. In all the interviews there are only seven such comments and the pupils usually responded differently in later interviews. They are nevertheless important because they represent some of the barriers to learning in the subject: a perception that one has to be 'religious' (Rudge 1998); an unwillingness to engage with 'exotic' beliefs; a lack of sense of what the learning outcomes are. The last point particularly highlights the need to share criteria, discussed above (section 8.3.3), but is relevant here because it reflects

pupils' concerns with the nature of religious education specifically. Overall, however these negative views highlight the importance of what follows: if some pupils sometimes wanted to reject the subject's educational rationale, the majority did not, and indeed were clear that they thought they could learn and did learn.

9.2 Pupils' views of 'learning about religion'

Knowledge and understanding of different beliefs is a generally regarded as a major part of religious education (though see Erricker and Erricker 2000). It is enshrined in the Framework as 'learning about religion'. As discussed in chapter three, there are arguments about its precise form, in particular the nature of 'religion' and 'religions', and what Jackson (1997) calls the issue of 'representation'. This section analyses both how pupils conceptualised this learning outcome, and how they described their progress towards it, but it will also consider this question of representation.

Pupils unsurprisingly perhaps envisaged understanding different religious beliefs as a learning outcome. However, the pupils almost never talked about 'religion' in the abstract. Over half either commented on beliefs and religions, or focussed on the people who held the beliefs and belonged to the religious tradition:

> Mike: understand other religions, what it's like to be in that religion, I think, the rules that you have to follow, things like that. [B3]

> Sara: just to learn about different cultures other than yours and your friend's cultures and to learn a bit about it and what - special occasions

and stuff. I quite like learning about that. [C2]

Importantly, though not forgetting those who did not see the subject's relevance, many saw this as valuable in itself and relevant to life outside school:

NF: Um, OK. Um, so what do you think RS is about?...

Ahmed: other people's >Annie: different religion< cultures

NF: it's a cultural thing? And what do think, why do you think people make you do it?

Annie: to understand other people...and that's important. [D1]

Dave: RE always pops up in day-to-day – activities, sort of thing. If you are chatting to your mates on say the London bombings and stuff, you have your/ know you have been taught stuff about Muslims, Hindus and stuff and you say to them about that and that's how they think, and like 'oh, OK'. [D3]

One pupil also identified limits to understanding, and the insider/outsider issue (see Arweck and Stringer 2002):

Kevin: yeah but you are never going to <u>understand</u> a religion, even if you actually joined it. [D2]

In general, almost all pupils thought that they could improve their knowledge and understanding in religious education. This is where one would expect the use of assessment for learning and self-assessment to be most easily developed.

Ahmed: If we did like um the Holocaust, I didn't know um, I knew briefly what it was but I didn't know in detail, but we did what happened and why, and now I have got more understanding of what happened, and other (*) and stuff. [D3]

NF: Do you think you've improved at RS?...

Aaron: yeah, definitely. 'Cause I was really rubbish, because I didn't know what Paul was, who Paul was and all the religions and stuff, but now I kind of like background and stuff so, I've got better at that [C1]

These examples show different pupils articulating a sense of progress in their knowledge and understanding, in particular topics. Some pupils emphasised the process of understanding other *people*, rather than simply an understanding of other beliefs or events:

Aaron: like understand their kind of views, because when I first came in I was like, well it's a hijab, kind of thing, you know. But now then, you can understand the views like from their point of view and from other people's point of view as well, which is much easier to learn and stuff. [C3]

NF: Chas, what about you then, what do you think you've learnt from RE? What have you gained over the last [term]?

Chas: like you say understanding other people's religions and stuff like that, because obviously (*) country and you know. But all them other ones and the different religions, it's Pakis and that, you know [A2]

Aaron possibly hints at a more empathetic view of understanding. Chas links

his understanding to cultural and religious diversity. It is not clear if he thinks this diversity is a good thing; though his use of 'Paki' is probably not intended to be derogatory, but reflects a limited vocabulary.

These pupils viewed knowledge and understanding as a potential learning outcome in religious education. Furthermore, they could make judgements about their own progress in them. Strikingly however, these pupils generally perceived this as being about the religious beliefs of particular groups of people, and not about 'religion' as an idealist abstraction. This supports Jackson's argument, particularly with Wright, that religious education should not be based on bounded reified abstractions such as 'Islam' or 'Hinduism' or 'Religion' but should use more fluid categories (see above, sections 3.1.1 and 2.1.2, and Jackson 1997, Wright 2004, Jackson 2008, Wright 2008). This is not to deny the usefulness of these general terms to describe general patterns, and indeed the pupils could also talk about Muslims and Islam equally, but it does suggest that they approach the study at the human level. This last point however needs to be treated with some caution both because the lessons and modules were conceived within a broadly interpretive approach; I never asked them to comment on 'Religion'.

9.3 Pupils' views of learning from religion

In this section, the data on the various reflective processes are considered. Three categories emerged. Since I was applying a broadly interpretive approach, the first two are labelled using Jackson's first two categories of reflexivity: 'constructive criticism' of the views of those studied; 'edification' of one's own views through reflection on them (described in section 3.2). They roughly matched Grimmitt's terms, Impersonal and Personal Evaluation

(Grimmitt 1987, 227-228). Wright describes similar reflective processes in the engagement of the horizons of the pupil and religion (2000, 181). However, a further category is presented: expression of pupils' own views.

9.3.1 Constructive criticism

The pupils identified the process of evaluating different beliefs as a learning outcome in the subject. As outlined in section 2.2, there is a discussion about the various elements of 'learning about religion', but the ability to comment personally on the material studied is key to many pedagogies, i.e. what Jackson calls constructive criticism (2004b), Grimmitt 'impersonal evaluation' (Grimmitt 1987). In terms of teaching, constructive criticism was an important part of the all modules, often linked to discussion.

> David: well you know what the beliefs are of the Christian. You are obviously going to be able to put a better opinion across, whether you agree with them or not. [A3]

> Aaron: yeah, you can see their kind of point of view of things, and if you don't agree with it then you can put your hand up and say 'I don't agree with that'. [C3]

Furthermore, pupils could self-assess their progress in this skill.

> Billy: yeah, I think I've improved, I can evaluate problems better, and um, answer questions about God, and, and it's all, it was going a lot better than last term I feel. [D1]

Dave: I find expressing my own views on other religions has kind of got better. [A3]

Although Dave talks of expression, he means a process clearly includes evaluation of other beliefs.

One pupil in particular had a clear concept of how her judgements on the beliefs and practices of others had changed.

Helen: oh the other night, I saw there was um an Indian tribe who sacrifice the third baby girl and in year 8 um, in year 8 I would have thought that was really harsh, but now I kind of understand because they have their beliefs and we have ours. [A2]

She considered that she had moved from a judgmental, absolutist approach to evaluation to a more qualified relativism on Punjabi female infanticide. This is certainly not to argue that religious education is *therefore* about engendering relativism. However, it is to identify this as part of one pupil's trajectory. Unfortunately, the issue was not pursued in the interview: what did Helen mean by 'kind of understand'? Is she now arguing for relativism or not? In any event, it would be a mistake to assume that this was the end of her development in this respect; indeed her thinking may have changed since then. This example is striking in another way too. She has applied her ability to evaluate to material that she encountered outside the classroom. She considers that her religious education has taught her to be less absolute in her judgements when confronted by challenging content both in and out of school.

To summarise, most pupils could identify evaluation as a learning outcome in

religious education, and they could also self-assess their progress in it. They understood it as contributing towards their understanding of the material they had studied, towards their ability to reflect on this material, and on their ability to reflect on religious beliefs that they encountered elsewhere.

9.3.2 Edification

If constructive criticism is one part of the process, edification is too, as discussed in the literature review. Almost all pupils were positive about this as a possible learning outcome and they could identify their progress towards it:

> Dave: how religious people have their different view on life, as, and how you can reflect on how your life is to theirs, um. It's quite good to know. [A3]

They also felt that they were better at it. One particular metaphor was that of putting both sides:

> Dave: I have learnt how to argue both sides of it, of the whole aspect of the subject. [A1]

> Kevin: I think I got a bit better, yeah. I can start to see two sides of a story. In year 7, it was like Ahmed said (*just one?) but now it's been showing two sides of the story. [D1]

They identified the value of being able to reflect on new ideas:

> Jake: so you are not just focussed on one religion that's yours or the

one that you came up as, and just kind of look at other things and don't just accept your one. [B3]

Further, they related this to knowledge. Thus they could see how a better understanding of beliefs would refine their opinions of them.

David: well, you know what the beliefs are of the Christian. You are obviously going to be able to put a better opinion across, whether you agree with them or not. [A3]

Importantly, they also recognised that their own positions could change as a result of this process, rather than simply expressing their views more effectively.

NF: does it make your views stronger? Does it make them more thought out?

Kevin: well, er it makes you think about them, which can, which can make them stronger or it can change them around. So that's stronger. [D3]

This was most striking when they commented on their progress in the module on the banning of *hijab* in France and Britain. They highlighted the different ways that they could react. It could strengthen their views:

NF:...What about, what about your own views, your own views on - respect for other people or – I mean the hijab thing is another example of that, you know...Do you think it changes people's views on that kind of thing?

Ahmed: yeah

Kevin: I think if anything it has made me more against it. [D3]

What is unfortunately not clear from this is whether and how Ahmed's views have changed. Is he commenting on changes to other people's views or his own? Other pupils also recognised the possibility of a change of mind as a result of their studies:

> David: you understand their point, but it could either make me more tolerant or it could make you more against it, it's just completely, it's up to you isn't it? It's up to you. It's your opinion. It helps your opinion no matter what. [A3]

What David suggests is that a deeper understanding of the reasons for wearing *hijab* has strengthened his view that women generally should not have to wear *hijab*. This is not the same as saying that *hijab* should be banned. However, he highlights the way that it 'helps your opinion'. In terms of the issues raised at the start of this section, we can see that he understood edification both as separate from constructive criticism but also as sometimes inseparable from it, as both Grimmitt and Jackson suggest. To decide to be against something is both to consider its relevance to ones' own point of view, and to evaluate its place in the lives of others.

9.3.3 Expression

One theme that emerged very strongly from the data was the value and importance for the pupils of expressing their own beliefs. This has an important place in many current pedagogical approaches. The current national guidelines suggest that at Key Stage 3 (11-14 year olds), pupils should 'express their own

beliefs and ideas, using a variety of forms of expression' (QCA 2004, 28).

The importance of this is articulated particularly strongly in dialogical approaches to religious education, which take pupils' own expressions of their beliefs as the starting point for the subject. Ipgrave for instance writes that 'as pupils are encouraged to contribute their own understanding...public recognition is given to the importance of that personal history (Ipgrave 2001, 14). This was important for almost all pupils; for example:

> Helen: at school, this is so much better than my old primary >Chas: yeah< school, we can do such more in school, because in my old primary school we weren't even allowed to say our own beliefs. We were told off if we said our own beliefs. [A1]

Many pupils particularly felt that they had improved at expressing themselves, particularly in discussions:

> NF: what about you? What do you think you have got better...at?
> David: definitely express myself now...definitely. [A3]

> NF: OK so what do you probably think you got better at? Listening to other people, at discussing >Dave: yeah< and understanding what other people think
> Chas: talking to people and putting my points across...
> Dave: expressing my points. [A2]

Just as evaluation is often linked to reflection, so also expression of one's own beliefs is often linked both to reflection and evaluation. This can be seen in the

remarks in the previous section about 'two sides' and it is particularly the case in discussion. A discussion cannot simply consist of different pupils expressing themselves; it requires a dialogue of comment, agreement and disagreement, and this enabled pupils to formulate their own ideas.

> NF: so how do discussions help you do that?...
>
> Jake: you learn about your own big topic through other people's opinions. Different ways.
>
> NF: so the way into the topic is through your beliefs?
>
> Ellie: yes
>
> Jake: and then you just have to have your opinion and then you get the whole picture. [B2]

These pupils highlight how they learn about a topic by engaging with the views of those around them. Thus there is not a sequential process of learning about religion and *then* learning from it. As Jake suggests, pupils learn holistically, from having 'your opinion' and the 'whole picture'.

9.3.4 Implications for the current attainment targets

This research suggests that although the two elements of the current attainment targets are important, they do not do full justice to the ways that pupils can make sense of their own learning, discussed below in section 9.5. It supports demands for one integrated attainment target for all elements. For pupils, the development of their own views was not separate from their understanding of the issues raised in the material that they explored.

> Dave: how religious people have their different view on life, and how you

can reflect on how your life is to theirs. It's quite good to know. [A3]

This avoids the problem that OFSTED identifies:

> Separating assessment often leads to tasks which invite pupils to offer personal responses which cannot be judged in terms of any real achievement or which tests them to recall or recount information about religion (OFSTED 2007, 39).

This supports demands for one integrated attainment target for both elements; Copley suggests that pupils should be 'engaging with religions and other life stances' and religious education should 'induct children in an informed way into what is a debate, extend their options, dispel ignorance, reveal hidden assumptions' (Copley 2005, 128). For instance, the process can encourage pupils to alter their views:

> Kevin: well, it makes you think about [your views], which can make them stronger or it can change them around. [D3]

Kevin has become both aware of more options, and of his 'unquestioning presuppositions are challenged' (Jackson 2000b, 136). At the very least, therefore, if pupils can synthesize the two intellectual elements in religious education, it would be sensible if pedagogy reflected this.

9.4 Self-assessment of the development of values

As has been shown, one of the main rationales for religious education in a pluralistic society is its contribution to civic values (section 3.3). But do pupils

think that they have developed these values as a result of their learning in religious education? Pupils should be able to self-assess their progress in all aspects of a subject, but if they cannot do this then it raises doubts about the appropriateness of such a rationale for the subject.

Before discussing the data on values, it is worthwhile recalling those pupils who thought that the subject had no merit. Not all the pupils had positive things to say, and some were silent in the discussions on these issues. However, many did consider that the subject had wider outcomes than the academic and that they had made progress in these.

9.4.1 Tolerance

Pupils recognised that this was a possible outcome of the subject. All who commented talked of tolerance to people, rather than to a religion, and they also explicitly linked it to reducing violence:

> Alan: the purpose of RE is tolerance...to tolerate other people, otherwise you wouldn't know about it would you? [C2]

Alan links tolerance to knowledge. Incidentally, his remark is all the more striking, given that he had earlier thought that religious education was 'crap'. Penny explained the connections more fully.

> Penny: yeah, I think it's good to know you/ about different religion's views and things. Because people can be very like -
>
> Dave: narrow minded
>
> Penny: narrow yeah, very narrow-minded against other religions, maybe

because they have a slightly different belief to them - and I think that they are quite, it can sometimes turn into like um maybe racism and it can=

=David: conflict between two religions

Penny: yeah, it can turn into quite maybe sometimes quite a nasty conflict if someone is like – very ignorant about someone's beliefs >David: people aren't aware of them< yeah and it can also like quite upsetting for maybe someone in the class who, like, has different beliefs >David: and believes them, yeah< and like different religions can probably be quite upsetting for somebody to sit there in a discussion saying 'No, I don't believe that, I think that's a load of rubbish. Your family are stupid for thinking that >David: umm< um, you, you shouldn't have those beliefs'. And that can probably be quite upsetting. So, I think RE teaches...people to be more tolerant. [A3]

Penny's comments link the individual's beliefs both to classroom dynamics and their home background. She sees the connection between classroom learning and a wider sense of identity and the values associated with it.

Further a handful also recognised that they had become more tolerant; they could self-assess their development in this disposition.

Kevin: I admit, I admit I could be, a bit, er -

Annie: pig-headed?

Kevin: yeah - um, but doing discussions has made me think that it isn't, it's about people's (*accounts?), as well, so I'm less intolerant [D1]

Here he highlights the activity, but he also commented on the work on the

because they have a slightly different belief to them - and I think that they are quite, it can sometimes turn into like um maybe racism and it can=

=David: conflict between two religions

Penny: yeah, it can turn into quite maybe sometimes quite a nasty conflict if someone is like – very ignorant about someone's beliefs >David: people aren't aware of them< yeah and it can also like quite upsetting for maybe someone in the class who, like, has different beliefs >David: and believes them, yeah< and like different religions can probably be quite upsetting for somebody to sit there in a discussion saying 'No, I don't believe that, I think that's a load of rubbish. Your family are stupid for thinking that >David: umm< um, you, you shouldn't have those beliefs'. And that can probably be quite upsetting. So, I think RE teaches…people to be more tolerant. [A3]

Penny's comments link the individual's beliefs both to classroom dynamics and their home background. She sees the connection between classroom learning and a wider sense of identity and the values associated with it.

Further a handful also recognised that they had become more tolerant; they could self-assess their development in this disposition.

Kevin: I admit, I admit I could be, a bit, er -

Annie: pig-headed?

Kevin: yeah - um, but doing discussions has made me think that it isn't, it's about people's (*accounts?), as well, so I'm less intolerant [D1]

Here he highlights the activity, but he also commented on the work on the

holocaust in this respect.

> Kevin:...before you've studied about it and you didn't know 100% what went on, but now, let's say, you've got the main outline of what did actually happen, and the evil has become more upsetting, for example. [D2]

There is obviously a wider discussion about holocaust education that it is not possible to explore fully (see Salmons 2003, Maitles, Cowan and Butler 2006), but what is important is the focus on how 'upsetting' can be positive. There were however implications if pupils did not learn fully.

> Alan: if you find it hard, you are less tolerant of the religion because you don't understand it, whereas if it is easier then you are more tolerant because you know more about it, which is what RE is about. [C3]

If pupils found the subject 'hard', then they would not develop the values that lay at its core. However, there was an important qualification to this. The fact of studying about it did not automatically cause one to be tolerant, even if some knowledge was better than none: as Jake observed "you could know a lot about Judaism, but just not like Jews". [B1]. He supports Jackson's assertion that 'it is a mistake to assume that understanding and knowledge necessarily foster tolerance' (Jackson 2005, 11).

Nevertheless, the fact of studying it could merely foster a rather specific tolerance;

> Mike: I'd say I have become more tolerant to the Jewish people who

were affected by it so if, I don't know, their parents had been killed or sons or daughters...I'd say...that I have become more tolerant towards the Jews, but I am still no more tolerant to the Germans than I was before. [B2]

Thus, although the responsibility to be tolerant is often understood by adults to arise out of a universal ethic, pupils do not necessarily understand it in this light. This particular view echoes a wider concern that the teaching of German history in British schools implicitly encourages stereotyping all Germans as Nazis (BBC 2002). However, what is important is this pupil's sense of how the content had impacted on his judgement; it makes a difference.

A similar issue emerged in discussions about the module on *hijab*. Pupils were reluctant to say that it had made them more tolerant, because they did not accept that it was right to wear it, but they accepted that they understood *why* Muslim women wore it.

> NF: if you know, if you, let's say you, look, we did the stuff on *hijab*, does that - does that mean you are more tolerant towards Muslim women?
> David: you understand them more
> Penny: yeah, I think you understand the=
> David: you understand their point, but it could either make me more tolerant or it could make you more against it, it's just completely, it's up to you isn't it? It's up to you. It's your opinion...it helps your opinion no matter what. [A2]

David here picks out a distinction between being tolerant of *hijab*, understanding the women who choose to wear it, and formulating his own view.

This is discussed in relation to 'learning from religion' in section 9.3.2; his studies gave him a deeper understanding of the issues, and therefore the beliefs of those who thought it was appropriate. What is highlighted here is his recognition that this depth could increase his opposition to the practice. He illuminates Gardner's suggestion that there 'is no incompatibility between believing someone to be mistaken and being tolerant towards them and their ideas' (Gardner 1992, 81).

Overall then, one can discern three features of pupils' views on tolerance. Firstly, they identified it as a learning outcome. Secondly, they could self-assess their development in it, and lastly they could identify some limits to it.

9.4.2 Respect

Tolerance however can be seen as merely putting up with a difference of opinion, whereas respect can be seen as a deeper level of acceptance of difference. Milot for instance writes that:

> Respect of the Other in their difference goes further, in terms of training, than knowledge of this difference. It is a question not only of 'knowing' that the Other does not necessarily share our own beliefs or values but, above all, developing respectful attitude towards them that allows them to feel accepted in their identity[1] (Milot 2005, 18).

---

[1] Translation NF: Le respect d'autrui dans sa différence va donc plus loin, en termes d'apprentissage, que la connaissance de cette différence. Il s'agit non seulement de «savoir» que l'autre ne partage pas nécessairement nos propres croyances ou valeurs mais, surtout, de développer une attitude respectueuse à son égard qui lui

Many saw respect as a distinct outcome, beyond knowledge or skills. One pupil made the point forcibly, over two interviews:

> Alan: I say it would tie in with a multi-cultural society, 'cause otherwise um it wouldn't actually work if you just tolerated them, you'd have to understand them, and know them and respect them. [C2]

> Alan: I don't know, I wouldn't call it tolerance. 'Cause if you are saying 'tolerate them', it means you don't like them.
>
> NF: respect, then
>
> Alan: yeah. Respect. > Sara and Aaron: yeah< That's a better word
>
> NF: tolerate things you don't really like =
>
> =Alan: yeah. If you tolerate things, you put up with them. [C3]

He makes a clear contrast between tolerance as a basic type of acceptance and respect, which requires a deeper understanding and knowledge. But is this consistent with his earlier claim about understanding not being enough of itself to engender tolerance? It is, because there is an important difference in two types of understanding. In the discussion on tolerance he talks about understanding the *religion*, in the second, he talks about understanding *them*. The first suggests a focus on the system of beliefs, the second on the people themselves. It is the second that leads to respect. The distinction is easier to make in French between 'savoir' – to know facts, and 'connaître' – to know people; indeed 'savoir' is precisely what Milot uses to describe the basic level of tolerance.

---

permette de se sentir accepté dans son identité.

Other pupils highlighted respect as being more than knowledge of the difference between cultures.

> Ahmed: I don't think religions, skills, is just about them. It is about like having respect for other countries, so
>
> Billie: it sums it all up, really
>
> Ahmed: yeah. [D3]

This example also highlights the idea of appropriate behaviour in different cultures, which other pupils identified,

> Lucy: maybe if you wanted to travel – and say if you went to – Pakistan or something, then you would um know like about um their religion and you wouldn't have to question them, things like that. [B3]

> Jake: if there was a big festival in the street you wouldn't like be totally clueless and - you'd see what you'd learnt. It was Divali or something. [B3]

They particularly identified knowing how to avoid offending people in this context

> Aaron: how not to offend them, like if you are like Christian and they're Hindu or something and you say a joke but it's offensive to them. It's nice to know when you've crossed a line. [C2]

> Kevin: it's like the HSBC advert [*a series of television advertisements for a bank showing business people making social blunders in different*

cultures]. Because they say, that they understand different people...it's like that. You do it so you can understand the beliefs around you. So you don't do stuff that'll upset them. Like in the advert there is the one who puts his feet on the table and it's, it's a real insult. [D3]

They also felt that they could make progress in it.

> NF:...What have you learnt from studying RE?
>
> Helen: that there's different opinions, and we should all respect them. We should all respect other people's opinions. Right
>
> Dave: many people have different ideas and you should listen to them and take them in. [A1]

Dave's addition also illustrates Milot's point. It is more than noting that people have different ideas, it is more than just listening to them, it is also 'taking them in'. There is also perhaps a difference between respecting people and respecting their opinions. Helen however developed this idea in a later interview, when identifying learning outcomes:

> Helen: you see, I've lived in many places before I came here, so I would respect a lot of religions, and I came with that. What I need to improve on is um, whether I respect other people's opinions, and the way that if um they get upset when I say my opinion, I will understand. [A2]

For her, it was not simply 'respecting' the religion, but being sensitive to how other people could react to her views. She did not feel that she should not express her views, but that she needed to appreciate how others would react.

This can be seen in the notions of 'relating' or 'talking' to different people.

> David: I'd like to be able to respect and talk to anyone that I want to in the street...know whatever their religion was, so I could talk to anyone about Hinduism, to a Jew about Judaism. [A2]

Pupils often highlighted the value of discussion in achieving this, which takes us full circle to the start of this chapter, since discussion was one of the central parts of the subject.

> Ahmed: discussion, and discussions, in classes. In practice the one we do, it was quite good, and so you get different opinions, um yeah. When you understand what other people think and different points of view
> Kevin: and also you are teaching (*to respect)
> Ahmed: yeah. [D3]

Overall, the pupils' remarks on respect reveal a number of themes. Firstly, many showed a broad awareness of a difference between respect and tolerance, even if they did not all articulate it in this way. Secondly, they could assess their own development towards it, and indeed identify complex patterns of sensitivity. Thirdly, they linked this to situations outside school, particularly the ideas of being able to understand and 'talk to' other people.

9.5 Progress in religious education

This chapter addressed the second research question, with its two sub-questions:

- How do pupils consider that self-assessment helps their progress in religious education?

  a. What do pupils consider to be progress in religious education?

  b. Do pupils think that they improve at this?

While a few pupils did not recognise the value of religious education, the second section showed that almost all were able to identify intellectual skills as learning outcomes, and also how they had made progress towards them. The third section showed that many pupils are also able to identify both a range of values as learning outcomes, and how they had made progress towards these.

There are two issues that arise from this, the contrast in the two different types of learning outcome, and the relationship between the content, activities and the types of learning outcome. Firstly, it is particularly striking that pupils were able to vocalise, justify and discuss the way that their values had developed as a result of the subject. These elements are essential justification for much current pedagogy of religious education (see chapter three). However, they are often neglected at the expense of the cognitive learning outcomes in attainment targets, as in the Guidelines and the Framework. This is not to suggest that the cognitive outcomes should not be developed, or are less important. But what this evidence clearly shows is that pedagogies of assessment in religious education need to take account of pupil development in this respect. In other words, an approach to assessment that simply tackles what is more easily measurable does not do justice to the potential range of intellectual and moral development that the subject demands. This is a serious shortcoming in the Framework; while it recognises the importance of attitudes and values, it omits to explain how they might be assessed because it is constructed in a constricting readerly summative paradigm.

Secondly, although the different elements in the learning process have been teased apart in presenting the data, pupils often thought of the process as a whole. It was the facts about the holocaust or the wearing of hijab to which they automatically reacted. Discussions on the issues were a way to understand, evaluate, reflect and become more tolerant. Thus, there was no major separation between knowing and evaluating, or between reflecting and respecting. For example, with my annotations:

> NF:...What have you learnt from studying RE?
>
> Helen: that there's [1] different opinions, and [2] we should all respect them. [3] We should all respect other people's opinions. Right
>
> Dave: [4] many people have different ideas and [5] you should listen to them and [6] take them in [A1]

This describes the subject as [1] knowledge and understanding, [2] an obligation to respect opinions, [3] and respect people, [4] recognition of difference, [5] respect of other beliefs through listening, and [6] evaluation and reflection. Between them, these two pupils map out a wide pedagogy in four lines. There are arguably other aspects of the subject that they have omitted – the sense of critique or evaluation is not developed. However, one needs to bear in mind that they are describing the areas where they feel that they have made progress, not just what the subject is. It is all the more striking that their sense of their own progress is seamless; they do not separate out the different elements. This complexity calls into question common presuppositions about a hierarchy of intellectual skills (Bloom 1956)

Finally, self-assessment has played a major part in this progression. Apart from

the data analysed in the previous chapter, what is striking is the way that pupils highlight discussion as the paramount vehicle for the self-assessment of all kinds of learning.

> Kevin: yeah - um, but doing discussions has made me think that it isn't, it's about people's (*accounts?), as well, so I'm less intolerant. [D1]

> Jake: you learn about your own big topic through other people's opinions. Different ways. [B2]

Further, the interviews themselves became a vehicle for further discussion and reflection. In a number of interviews the way that other pupils comment on, develop and embellish remarks and ideas further suggests that they found the process rewarding. In some respects this picks up a concern in the analysis of the data that there was an evolving collective view, which was difficult to ascribe to a single pupil (see section 6.5.3), but what was a concern from the research perspective is possibly valuable evidence of the reflective potential in pupils.

9.6 Conclusion

This chapter is both significant and distinctive in addressing the academic issue of the pedagogical nature and status of religious education, notably the inter-relationship between the intellectual and ethical dimensions of learning in religious education, as well as the practical issue of how pupils can understand their learning in the classroom.

Through the empirical research, pupils' views of their learning in religious education have been analysed, both in terms of their understanding of the

current attainment targets, and their sense of progression towards these targets. This has shown that pupils are able to identify: progress in their knowledge and understanding; progress in their ability to evaluate their ideas of others; progress in how they reflect on and express their own views; and subtle changes in their attitudes and values. Moreover, it shows how pupils understand these three processes as inter-related, which is not how they are currently presented in the Framework.

However, while the previous two chapters have contributed to this research by addressing the two fields of literature separately, the next chapter will combine them. It analyses how three pupils' views of assessment processes are combined with their views on religious education in order to make more sense of the inter-relationship of assessment and pedagogy in pupils' classroom experience.

Chapter Ten

Self-assessment in religious education: three pupils' voices

So far, this research has consisted of two separate strands, one on assessment for learning and the other on religious education. There are two reasons for this. Firstly, the two fields of literature have almost never overlapped and secondly because the pupils themselves rarely discussed the two processes together. The analysis so far has followed this distinction, and has revealed the constituent elements in the light of these separate fields.

In broad terms, each strand has revealed particular issues. In the analysis of assessment, there is a tension between ego-focused and task-focused perspectives in pupils (Dweck 1999), and their reactions to self-assessment are coloured by this. In the analysis of pupils' understanding of their progress in religious education, there is a subtle interrelationship between, broadly, learning about religion, learning from religion and attitudes and values, supporting arguments for a pedagogy that combines these elements, notably an interpretive approach (Jackson 2004b).

What is now required is a synthesized approach, but where and how do they inter-relate? Even though the pupils rarely discussed the areas together, *all* the pupils talked about both of them in some way. The fields therefore are linked in the educational experiences of the pupils. Thus, this chapter will consider the views of particular pupils to identify the connections between their attitudes in both areas. It is to a limited extent biographical, in that it focuses on the

individuals and moreover can be seen seeking 'interpretive validity', which means catching the meaning, interpretations, terms and intentions that situations and events have for the participants, rather than terms of the theoretical issues (Maxwell 1992, 290).

10.1 Pupils' experiences of assessment in religious education

Recent research has highlighted the need to examine the pupils' experiences in school as a whole in order to make sense of their attitudes to assessment. For instance, Filer and Pollard examine one child's experiences of assessment throughout primary school (2000, 36-52). Ecclestone and Pryor use the term 'assessment career' (see section 2.5.6), as the subjective longitudinal dynamic of pupils' experiences of assessment. In particular, they ask whether this notion can 'help teachers improve the impact of formative assessment on motivation and autonomy for learning?' (Ecclestone and Pryor 2003, 485, see also Ecclestone 2007).

Three pupils' comments on various aspects of their experiences of assessment in religious education will be considered as moments in their longer assessment careers. The three pupils have been selected because they represent a high, medium and low attaining pupil, judged by the general data that I had as a teacher, which were their SATS results at the end of primary school and CATS tests at the start of secondary school. Both sexes are represented, though the issue of gender is not explored. They also represent differing positions in terms of whether they are ego-focused or task-focused (see section 2.2 above, Gipps

1994, Dweck 1999), though in more complex ways than a simple dichotomy of types. Some of their remarks were included in the previous chapters.

10.2 Aaron

Thin and freckly, Aaron had struggled throughout his learning career so far; indeed, he was identified as having specific learning difficulties. From a class management point of view, Aaron was not disruptive, although he was sometimes distracted. The guidance from the school's learning support department showed that when he had entered the school aged eleven, two years before the research started, his reading age had been identified through literacy tests as being that of a notional average pupil of seven years exactly, and his spelling age that of seven years and eleven months, i.e. significantly under the norm in both. These were the lowest in the class, in which others had had the notional reading and spelling ages of average fifteen year olds; put starkly, according to the data, other pupils were over twice his reading age. In his SATS tests at the end of primary school he had achieved level three in maths, science and English (the norm is level four); others had achieved level five in them all. By the start of the research, his reading age was identified as being nine years and a month, and his spelling age as being eight years and two months. He had therefore made steady progress in his reading, and less in his spelling, but was still considerably less developed in these skills than some of his peers had been two years before. More detailed school comments stated:

> Easily distracted; responds well to help and praise; written work is slow and laborious; provide clear examples of what he is aiming for &/or writing

frames; set time restricted tasks; tends to sit for ages doing very little otherwise; provide key vocabulary for him to learn in small amounts.

These comments showed that other teachers had identified successful formative strategies, such as sharing criteria through clear examples, and also he had a disposition to accept help.

From the perspective of research, Aaron was the type of low attainer who was at risk of under-achieving as a result of the experience of constant comparative failure in tests (Harlen and Deakin-Crick 2002). However he was also the sort of pupil who could benefit most from assessment for learning (Black *et al.* 2002, 9). Was he a task-focused pupil, worn down by failure in a competitive environment, or could he seize the opportunity to become more ego-focused?

It is perhaps no surprise that Aaron vehemently disliked summative assessment, for example when asked about being given levels on his work:

> NF: yeah? Um, those are all ways of assessing yourself. Um, some of them use levels
>
> Aaron: I don't like levels
>
> NF: sorry?
>
> Aaron: I don't like levels
>
> NF: You don't like levels
>
> Aaron: I don't like getting my levels, personally
>
> NF: OK
>
> Jenni: why not?

Aaron: I'm not just saying that. But, you know, I don't like levels

NF: OK, well, I mean that's interesting, that's one of the things I wanted to, to ask you about really was, was, was um – if you did like levels. I mean do you – you [to Aaron] don't like levels.

Aaron: No

NF: why not?

Aaron: 'cause I just don't like getting them, if they um, if they're too high I think "Oh yeah, I'm really good, I'm a boffin" kind of thing, and if I do too low I think "Oh I can do that better" and, and oh I just don't like levels, completely. I just don't, any subject, um

NF: is that, you don't like being given levels by the teacher or

Aaron: yeah, I just don't like being given levels by a teacher because I just don't like them. [C1]

The emphatic repetition of the phrase "I don't like' levels" is very insistent. It is clear that this is an emotional event. Over and above this litany, he could identify the processes at work in ego-focused attitudes: the dependency on the teacher's judgment; the attribution of attainment to internal qualities like intelligence ('boffin'); the demotivating effect of low levels. In the second interview, he explained further why he found getting levels demotivating, as Butler would suggest (Butler 1988).

Aaron: if you get a lower level you feel a bit, I dunno, oh God a bit down, like you'll get an even lower level...yeah it's a down, downward hill. [C2]

By contrast, not only was he enthusiastic about formative approaches and techniques, he was also able to explain why he preferred them: for instance, on feedback:

> NF: OK. Comments or levels, which are better from the teacher?
>
> Aaron : comments, um, yup. [C1]

> Aaron:...[if] you get a level you don't really know what to improve on because they don't actually write it down and say "you need on this that and the other". [C1]

He also liked the use of the lesson plenary to review progress:

> Aaron: yeah. It's all right just to know that you've done something in the lesson, and stuff so, and say if its good or bad...yeah, but if you achieve it you can just like tick it, and if you don't, it's just like a simple tick or cross, smiley face, sad face, something like that. [C2]

He thought that they were useful because 'you could get advice from them' [C1], and peer assessment 'was all right'. [C3].

From the start of the research, he was able to describe self-assessment in an accurate (if limited) way:

> Aaron: Yeah, assessing yourself as in what you know about what you've just done in the last week or so. [C1]

The limitation is that there is no sense of using the process to direct and control his learning, but simply as checking what has been covered. However, given that some of his peers had thought it was like a personality questionnaire in a magazine, his reply suggests some awareness of reflection.

Surprisingly perhaps, given his dislike of levels from the teacher, he thought that the use of levels in peer or self-assessment was beneficial. At the end of his tirade against levels from the teacher, he responded thus:

> Aaron: yeah, I just don't like being given levels by a teacher because I just don't like them
>
> NF: right. So when we did it that you had to work out what level you were, and what you wanted to aim for that, was that still bad or was that a better=
>
> =Aaron: that was OK, actually, that was quite good [C1]

This is surprising since one might imagine that he would be wary of the use of levels in any context.

Assessment for learning seemed to offer Aaron the chance to escape from the tyranny of marks and levels, but it also allowed him to use marks and levels formatively; this points to more complex processes than Butler's research might suggest (Butler 1988), in which the low-achieving pupils who were given both marks and comments tended to ignore the comments, and became demotivated by the marks. He suggests that he can use summative assessment formatively, under the right conditions. This shows how the challenge to achieve a more

positive relationship between formative and summative assessment can be met (e.g. Black *et al.* 2003, 55). He represents the kind of learner who had become disillusioned and disaffected by a summative assessment regime, but was still committed to education, and would respond in a task-focused way given the opportunity.

In tandem with his ability to benefit from formative processes, he was fully able to reflect on his improvement within the subject.

> NF:...Do you think you've improved at RS?...
>
> Aaron: yeah, definitely. 'Cause I was really rubbish, because I didn't know ...who [St] Paul was and all the religions and stuff, but now I kind of like background and stuff so, I've got better at that, um yeah, that's it [C1]
>
> Aaron: I think you can like understand their kind of views, because when I first came in I was like, well it's a *hijab*, kind of thing, you know. But now then, you can understand the views like from their point of view and from other people's point of view as well, which is much easier to learn and stuff. [C3]

He identified the problem he found in writing down ideas in a structured way:

> Aaron: well I think I could do a little bit more research really because – well, I did quite good in my essay thing but it was a bit sort of (*)... I could do more. I have all these kind of like ideas in my head but I just can't make them on to paper [C3]

He also had a view of the general educational goals of religious education:

> Aaron: Other people's religion. If you have a religion then – er – you can see other people's point of view and how their religion works, and stuff like that, and if you meet them you'd know how to not insult them. [C3]

> Aaron: I'm not really religious but, so you could pick up tips from other people's religion, which, I think, are true. Well, it's not a particular religion, but – it could be on things like life after death, sort of things and hell and stuff like that (*) so pick up (*and you think) 'yeah, well that's true' [C3]

What can be seen more generally however is that formative self-assessment strategies allowed Aaron to identify his progress in religious education without recourse to the use of levels all the time. He could conceptualise his learning, and describe improvements in his learning in a number of ways, from knowledge and understanding to reflection to a more general vision of the educational goals. This helped Aaron develop 'learning autonomy' (Black *et al.* 2006).

The area where assessment processes and religious education came together was in discussion.

> Aaron: yeah. Discussion does help like loads, I'd say because you get other people's points of view and sometimes you, I actually change my mind, plenty of times because of what other people have said, I go 'er that's interesting'[C2]

As explained in section 7.5.4, it was often difficult to tell if pupils were commenting on discussions generally or about this specific use.

> NF: Is discussion a good way to do that or should we do it by, I mean we could like work out a series of levels for it and we could say actually you have got to decide what level of tolerance you are
>
> Aaron: nerrr
>
> NF: is discussion better?
>
> Aaron: discussion is better> Alan and Sara: discussion's better < because you are listening and talking [C1]

Aaron seemed to reiterate a general point. In a later interview, I tried to get him to elaborate on this difference:

> Aaron; yeah, you can see their kind of point of view of things, and if you don't agree with it then you can put your hand up and say 'I don't agree with that'
>
> NF : the discussion we had, particularly that I had in mind was the one at the end of doing the holocaust, which was whether you have become more tolerant or not, rather than, so the discussion is not just about – what people think, but – there'd be, you know, how you think you had become more tolerant
>
> Aaron : I thought that was a good one actually.... I thought it got other people's view [C3]

It was therefore ultimately difficult in the interviews to separate out his reasons for liking discussion as a vehicle for self-assessment from his liking of it generally. This however may point to the more significant fact that the processes were so similar for Aaron as to be the same. The reflective, meta-cognitive dimensions of both self-assessment and religious education combined positively for Aaron in ways that allowed him to escape from the negative effects of the implicit labelling of marks and levels. His learning was becoming 'reflective, intentional, and collaborative' (Black *et al.* 2006, 126). This exemplifies the suggestion that students need to know that their voice will be heard if they are to develop the deep-seated motivation to succeed (Munns and Woodward 2006, Deakin Crick *et al.* 2007). Moreover, this echoes O'Grady's findings about empowerment of pupils in religious education (O'Grady 2003), and my suggestion that pupils value genuine dialogue with their teacher (Fancourt 2007).

10.3 Jenni

If Aaron was keen to find a way out of labelling, Jenni was resigned to it. As Perrenoud remarks (quoted in section 2.5 above):

> A number of pupils do not aspire to learn as much as possible, but are content to 'get by'...Formative assessment invariably presupposes a shift in this equilibrium point towards more school work, a serious attitude towards learning (Perrenoud 1991, 92).

Jenni was an example of a pupil who was not prepared, at this stage of her education, to shift the 'equilibrium point'. In terms of the literacy test data, when

she entered the school aged eleven years and two months she had a reading age of eleven years and two months, and spelling age of ten years and one month. In her end of primary SATS results, she attained level 4 in all three subjects. This data therefore suggested that she was an archetypal 'average' pupil by the end of her primary education; however her medical records indicated that she had suffered from mild depression more recently. This did not mean that she was socially isolated in the class; she had several friends, and could be chatty. Clearly however her general psychological disposition would affect her ability to transform her learning in a sustained way, but what she has to say reveals how her general disposition maintains and is maintained by her assessment career. Her lack of engagement in lessons was also reflected in the interviews, and her responses were broadly consistent across the three interviews.

At the start of the research, Jenni could not see the value of the subject, even though she thought that she may have made some progress in it. She agreed emphatically with Alan's criticism of the subject but could not give reasons.

> NF: Anyone else? Jenni?=
>
> =Alan: not really, um I don't see the point in RE.
>
> Jenni: <u>noooo</u>. [C1]

Jenni gave some indication that she had made some improvement, but attributed this, at the start, to biological factors beyond her volition:

> NF: so you think you are improving?

Jenni: yeah...it's just my brain is getting bigger, that's all [C1]

NF: you might not like it but you still might think you have got better at it.

Jenni: I suppose I am a bit better at it, but still. [C1]

Jenni; yeah, I am definitely clearer [about RE] than I was at the beginning of the year. [C3]

It is difficult to put much weight on this sense of progress, given the tentative nature of the remarks and as she neither demonstrates it, unlike Aaron who gave clear examples, nor considers it is the result of her efforts.

She, like Aaron, disliked the use of levels and found them demotivating. As a pupil whose general school performance was likely to be in the middle of the class, she had developed a sense of 'failure' from tests and marks.

NF: does having, does having the level motivate you?...

Jenni: no, because you just look at the level and think "oh well that's it". [C1]

Thus, she seems generally to assume that she would fail, and furthermore if she used levels for target-setting, she would fail to meet the target.

Jenni: it was like sort of like what Alan said, like if you give yourself a level 5 and like get a level 4...you get like lower, you feel like really down and stuff. [C1]

> Jenni: well technically, I think...that um [targets] are a bit like a waste of time because if you do, you just feel bad and I take no notice of them. [C2]

In this remark, she does not even explain how and why she would 'feel bad'; it is axiomatic for her that she would be unsuccessful. Jenni exemplifies an ego-focused student; she shows the 'helpless' pattern (Dweck 1999, 5) because she views any apparent failure as beyond her control ("oh well, that's it"). In other words, she has no desire to 'close the performance shortfall' (Wiliam 1998, 5).

However, whereas Aaron clearly considered that he had made progress through the use of formative assessment strategies, Jenni considered many of them to be pointless. Thus in relation to the use of traffic-lighting for the lesson or the module:

> Jenni: its just like a waste of time because um - some might not change, like you might not like have improved much and you just look and you feel like really bad because you know what you do at the beginning of term and you haven't really changed. [C1]

Possibility of failure haunted her even when there was no comparative element within the class; she remained 'helpless'. This is striking in comparison with Aaron, whom one might have imagined would be more likely to think in this way.

As a result, she had a clear view of her strategy in school; it was about 'getting by' in "any subject" [C1]:

Jenni:...It's just a waste of, well technically, it's just a waste of time. It's just, you don't really care what you've done in the lesson. It's just like the end of the lesson, the lesson is over so yeah, >Sara; you want to go< yeah, you want to go, so you don't care what you've done, so

NF: right. As long as you've turned up and done what you've been asked to do, that's fine

Jenni: yup

NF: to be asked to think about whether you've done it is

Jenni: out of order! [C2]

The teacher's role was simply to set tasks, and the pupils could then try to do as little as possible. It was not that she obediently carried out all the work set. As she comments, she would not know what was expected of her, and moreover, she would be unlikely to complete any tasks set.

NF: so for you it's like come in do the work, kind of

Jenni; yea, and just walk out when the bell rings so

NF: right, you do what you are asked to do and=

=Jenni; well no not really [C2]

My final suggestion that she would do what she was asked was also rejected. For her, the lessons were about finding a way of getting through the day, of doing the minimum. Not for her greater 'learning autonomy' and the concomitant challenge (Black *et al.* 2006). She 'reject[s] approaches emphasizing intellectual quality in favour of repetitive, non-challenging and educationally debilitating work' (McFadden and Munns 2002, 361).

Unsurprisingly, this negative view of work was linked to a judgmental attitude to the efforts both of others and of herself, because being ego-focused led to her understanding assessment in a norm-referenced way, and therefore as a competition, and 'a form of social comparison' (Gipps 1994, 41).

> NF: OK. Are you worried about what other people think of your levels, do you think?
>
> Jenni: No, I just laugh along if they laugh at me. [C1]

She also commented that another pupil's piece of work for peer assessment was 'rubbish' [C1], and 'it's funny listening to people arguing' [C1].

Further, whereas Aaron found discussions a positive way of learning, Jenni was concerned with her ability to speak in public and her perception of how others would be regarding her. The *social* aspect of the comparison dominated her attitude.

> Alan:...I think you should be assessed on what you say, on what you do in a discussion
>
> Jenni: No! Because then I would get really low scores [C2]

> Jenni:...You've got all these people looking at you and it's like 'what are you staring at?' [C1]

What for Aaron and Alan was a way out of comparison was for Jenni an even worse form of comparison; to speak in public was to be judged. Nevertheless,

she did accept that listening to discussions changed her mind, but even here she was unwilling to enter into the process.

> NF: what do you gain from the discussions?...the discussions don't give you knowledge
>
> Jenni; no, a bit like a better understanding, something like that, I dunno, something
>
> NF: so they help you think?
>
> Jenni: yeah. They help me think [C2]

> NF:...for some people it's the process of listening
>
> Jenni: I change my mind, but I can't be arsed to like put my hand up and say it. I can't be bothered. I just can't. I just can't be arsed. [C2]

In terms of her progress in religious education, Jenni had only a little to say. As quoted above, discussions helped her to think, and she thought that she was 'a little better', with a clearer idea of what 'RE was about'. There were therefore, over the year, some signs of change, but overall Jenni generally remained sceptical about adopting a new attitude to her learning in religious education.

It seems likely that her reluctance was partly due to a wider experience of assessment that spread across her learning in school; her remarks are not subject specific, but are about levels and marks generally. School, including religious education, is not for her. In retrospect, a question about whether she felt she succeeded elsewhere would have been interesting. It would also be unwise to ignore her mild depression as a factor in her negative view of herself,

but it is striking how it feeds into an assessment career that inculcates an ego-centered view of learning. Despite my attempts to offer her a new way of being a pupil, she was unwilling to take the risk.

10.4 Kevin

Kevin was relatively tall for his age, feisty and a keen sports player. In terms of general ability he was above average. When he entered the school aged 11 ½ years, he had a reading age of a twelve year old and a notional spelling age of 12 years and 5 months. Although he had not been entered for the end of primary school SATS, his cognitive ability test scores suggested that his attainment in mathematics and science at the end of year 9 was likely to be better than average (levels 6b and 6c respectively). There were no other comments about his educational progress, but it was noted that he had difficulties with another boy in the class with behavioural issues, largely as a result of standing up to the other pupil in bullying incidents.

As far as the research process was concerned, he was occasionally frustrated by some of my questions, and prepared to say so; for instance, at one point he felt that I was reading more into their responses than I should:

> Kevin: you turn it around so it makes us look, look better. No, you just turn it round somehow and it makes, I don't know F [D2]

However he did not withdraw from the interviews, and indeed continued to play an active part both in that interview and the next.

Kevin's attitude to summative assessment was ambivalent. He had broadly succeeded in terms of achievement under the existing paradigm, and research suggests that higher attaining pupils continue to attain highly under summative processes and remain motivated (Butler 1988). Thus, he often appeared to be an ego-focused pupil. He generally liked to know from the teacher what level his work would attain.

> Kevin: <u>you</u> should mark it. <u>You</u> should decide what <u>you</u> think it would be [D2]

> NF: Does [it] help when you get a level back?
> Kevin: yeah, so you know if you've underachieved or got what you wanted, or better...'cause if you haven't [got] it, you get a crappy level, then you're thinking 'I don't know what level I was aiming for'. But if you know what level you were aiming for, you know you got a bad one [D1]

He also believed that getting low levels or marks would motivate him to improve.

> Kevin: I think if you gave me a level 2, it would make me want to do better, because level 2 is pretty poor [D1]

'Petty poor' is of course comparative; he also claimed that this was 'a competition', which suggests it is against others, as Gipps suggested (Gipps 1994, 41). Thus he was, to some extent, succeeding for the 'wrong' reasons (Harlen and Deakin-Crick 2002). However, this response contrasts with Jenni's, for whom receiving low levels was demotivating.

Consonant with this norm-referenced approach was his difficulty in identifying quality in his work, and so one criterion he could identify was simply the time spent on it.

> Kevin; I find it hard...I get confused.
>
> NF: you get confused about what the different levels are?
>
> Kevin: with Maths, it's easy because you have to get a certain amount of marks, to get a certain level. With RE, you have to say a specific thing, which, if you've got half and half, half of it's in the text and the right answer, and half of it isn't, I don't know whether that would be worth marks, or not. [D1]

> Kevin: if it's a good piece of work, normally you can tell...because you've spent a lot of time and everything. [D1]

But not only was he using the levels for norm referencing, it was also about earning external praise from the teacher.

> Kevin; you almost want a good level to – make – the teacher pleased with you, well I do
>
> Others: *Gasp*
>
> Kevin: well I know it's a bit sad, but I think if I get a good grade, they'll be nice to me> others: *laughter*< if you get a crappy grade then -
>
> NF: they'll be really horrid?
>
> Kevin: yeah...

NF: OK, so that's, that's so that's important because of what the teacher thinks

Kevin: yea, yea...I got it from playing sport. If you do, if you do your job, when you are playing a match, if you do your bit for the team, the teacher will be pleased with you and he'll talk to you as if you're an adult. If you're playing football and in the game you do nothing – they'll be disappointed with you and they won't talk to you as they would have done if you'd played the game...I've sort of done it many times, very sad, but it works. [D1]

Alongside this ego-centered approach, the last comment reveals a personal sense of frustration with the process, but one that he feels constrained to follow: 'very sad, but it works'. In this context the gasp of others is striking; this may be an issue of peer group pressure in group interviews – that he had admitted to this. Adult approval is not very 'cool' for adolescents. I understood it as surprise that as confident a character as Kevin would admit to it. He also recognized that the process could be demotivating:

Kevin: [getting levels] can do one of two things. It can make you want to work harder

Annie: or it can make you feel really dumb

Kevin; it can make you feel dumb because you don't what to do any more work because you know it will be the same standard. [D2]

The last remark illustrates the helpless pattern again, like Jenni. This process was linked to his image among his peers. Thus he did not like the process of applying levels to his own work because of the social risks of getting it 'wrong'.

> Kevin: if you give yourself a level that's above...you might look a bit of an idiot...
>
> NF: so, if you, if it's below then you are making yourself look stupid, but if it is above you are making yourself look -
>
> Kevin: a big mouth. [D2]

This concern also led to him suggesting strategies to avoid having one's work used as an example for peer assessment.

> Kevin: it could also actually er make your work worse, because you would write a bad bit of work (*so that) you wouldn't show it to the class [D2]

An attempt by a teacher to introduce this strategy could fail because pupils interpreted it in the light of their previous experiences, similar to my experiences of plenary reviews.

He was however open to some formative strategies, being enthusiastic about getting feedback as comments and not levels:

> Kevin: if, you've got comments, then job done. [D1]

> Kevin: comments...because it helps you see what's good in comparison – but with a level, just to (*say this is a four, but) yours is a level three...Whereas in a comment like he has got that, but I haven't...but if I put that in, then it'll be better. [D2]

Nevertheless, he could be ambivalent about others in two ways, either because his reasons seemed to mix task focused thinking with ego-focused thinking, or because he gave different responses at different interviews. Firstly, he valued peer assessment:

> Kevin: if you look at someone else's, you get an idea of what they are writing...if they are getting good marks, better than you, you think right if they are doing it, if they are getting better marks that me, I'm doing something wrong - and they are doing something right...you kind of take bits from theirs over to yours...and then when you get sort of better. [D1]

The phrase 'getting better marks than me' may suggest that he is thinking competitively about the process, but he also seems to see it formatively, because he focuses on what he and the other pupil are 'doing'. He does not attribute it to their intelligence as a fixed concept, but as about performance of the task (Dweck 1999, 20). Thus he rationalizes these formative techniques within his existing assessment career; it is both ego-focused and task-focused at the same time.

As an example of the second form of ambivalence, he initially liked the use of the traffic-lighting system over the module (see appendix 1).

> Kevin: at the end, you say what you can and you can't do. I find that helps - because when I do it the second time - I can see quite a drastic improvement from what I understand, and I can put, like a tick or something there...but I didn't think I'd be able to, when I was doing it in the first place, which is quite good. [D1]

Later, however he was less impressed. Firstly, he said that he 'found them boring for some reason' [D3]. Secondly, he also had practical objections to the way that it had been carried out.

> Kevin: so Miss Gale [the science teacher] always sticks it in the book, because otherwise I'd have lost it. Because in science we do it. So she always sticks it in for us, so we can't lose it, sort of thing. So it's almost like, we look at it and think, 'I know what that is now'.
>
> NF: so you have done them in two subjects and – do they work in both, do they work better in one?
>
> Kevin: in science it has worked because she has taken in the book for us, but in RE, it didn't really work....It gets lost. Because we have got so many, and they are only half full, so which one do I need? [D3]

Thus for Kevin, the failure of this process was partly my failure to ensure that very simple practicalities were carried out. It seems ironic that the development of independent learners could require so basic a foundation.

By contrast, there were some activities that he was permanently opposed to throughout the interviews, notably the use of a rapporteur to feedback at the end of the lesson. This activity was abandoned quite early on in the research (see section 7.3.2), and Kevin's response was probably partly due to the fact that he was one of those very few rapporteurs at the front of the class; his experience was very personal.

> NF: ok, getting one person to do it. 'Cause we tried that briefly, we tried that, but it didn't quite happen. You wouldn't want to do that 'cause of the sort of embarrassment thing – you did it once didn't you Kevin? What did you think of it? Was that better or worse, or
>
> Kevin: I didn't like it...no, I don't like speaking to groups of people [D1]

> Kevin: You used to have to go up to the board which was (*embarrassing). I think never really got into it [D2]

It is striking that yet again the socio-dynamics of the class was the cause of concern for him, and indeed he recognized that if a different classroom atmosphere had been developed, then it could have been suitable, as the Learn how to Learn project suggested (Learn how to learn 2002).

> Kevin: if we had started it in Year 7, then we would have found it much easier, but the fact that it was started in Year 9, then it was a bit of a culture shock. [D3]

The phrase 'culture shock' is important too; he recognizes that a culture had developed in the class, which would be hard to change. To sum up his attitudes to formative approaches, Kevin was broadly prepared to experiment with them. He was not as enthusiastic as Aaron, because summative processes though 'sad' still 'worked', but he was more positive than Jenni.

Alongside this, he also reflected on his sense of progress in religious education. Kevin had a strong sense of progress in a range of skills.

> NF: the question was do you think you've got better?
>
> Kevin: um yeah, yeah, yeah [D2]

> NF: does it make your views stronger? Does it make them more thought out?
>
> Kevin: well, it makes you think about them, which can, which can make them stronger [D3]

However, he was also able to articulate a wider social understanding of the educational goals of the subject.

> Kevin: so that if you go to a different country and they study, worship a different God, or a different religion. [D1]

> Kevin: it's like the [bank] advert [*a series in which business executives are seen behaving inappropriately in different cultures*]. Because they say, that they understand different people, like in er is it Mexico?...It's like that,

> you do it so you can understand the beliefs around you. So you don't do stuff that'll upset them. Like in the advert, there is the one who puts his feet on the table and its, it's a real insult. [D3]

In particular, Kevin believed that a value of the subject was in developing "tolerance" [D3], and his learning had helped him to do this.

> Kevin:...doing discussions has made me think that it isn't, it's about people's (*accounts?), as well, so I'm less intolerant [D2]

> Kevin: it makes you much more inclined to listen to what [other people] have got to say [D2]

He also felt that this attitude had extended outside the classroom.

> Kevin: I believed [the module on *hijab*] helped. It made me feel more comfortable joining in a conversation about it 'cause I understand more about what's going on. So did my parents talking about it, instead of me thinking what the hell are they on about, I could join in. [D1]

In these quotations, Kevin identifies particular learning activities, e.g. discussions, as being the key to this development (see OFSTED 2005). Apart from the fact that he was able to make this connection, what is also striking is that while he was wary of many assessment processes that involved speaking out or putting his work 'on display', he was nevertheless more than prepared to contribute to and identify the benefits of speaking out in a different sort of

classroom process, and indeed to consider that the attributes that he developed are relevant both inside and outside the classroom. This can be compared with Jenni, for whom any classroom speech involved judgment by her peers, and who did not connect the attributes in the classroom to her life beyond school.

Kevin also recognized that this process would not be easy or automatic. In discussing the ability to empathize, he emphasized the importance of volition and patience.

> NF: right, to find out what it was like. OK. I mean a lot of people said um, in terms of RE generally there is a big discussion as to whether, really you can put yourself in someone else's shoes, whether that's really possible. Um, but you are saying that doing that helps you along a process
>
> Kevin: It depends how much work you put into it. If you just sit there thinking 'I don't want to do it' then you won't...You've got to sit there thinking 'I can do this and I want to do this', and as soon as you start thinking that it will start coming to you in time [D2]

In this respect, Kevin is very task-focused. He believes that it is effort and will that will enable him to succeed, and recognizes the need for patience. 'Work' in this context is not the written text but a mental activity.

Overall, Kevin presents a complex picture. He is both task-focused and ego-focused. On the one hand, he can negotiate his way through the competitive aspects of more summative processes, particularly as he is relatively successful

at them. On the other hand, he can appreciate the benefits of formative processes, such as peer assessment and self-assessment. Further, he particularly valued discussion both as a means of assessing and contributing towards the development of his attitudes and values. However, he accepts and then rejects different self-assessment techniques during the year.

He shows how a simple dichotomy between being ego-focused and task-focused may not do full justice to the patchwork of ideas and attitudes that pupils may have to their learning. He is both positive and negative about experiences of both assessment paradigms at the same time. Further, a practical problem he represents is how teachers can help high-attaining students to move from an ego-focused approach to a task-focused one, i.e. to abandon an approach in which they succeed.

10.5 Assessment careers - discussion

What do these glimpses into the minds of three Year 9 pupils reveal about pupils' perspectives on learning and assessment? There are five issues (from Ecclestone and Pryor 2003, 485): engagement with formative processes; motivation and autonomy; pedagogy and assessment; institutional issues; social and psychological issues.

Firstly, how do students engage with the formative assessment strategies offered to them? Importantly, the pupils engaged in a variety of ways – even in one class. Jenni scarcely engaged at all. Aaron and Kevin were certainly enthusiastic about some of the strategies at least some of the time, but in very

different ways. There was a tendency for all of them to interpret strategies such as peer assessment using levels that were intended to be formative in a summative way. The two boys were particularly enthusiastic about the use of discussion, and its contribution to their ability to reflect on the development of values. This clearly supports the idea that pupils are on a spectrum from ego-focused to task-focused, but they are also on a spectrum in terms of their receptivity for formative techniques, from Aaron's wholesale adoption, to Jenni's rejection. Further, however, it suggests that different pupils will find different strategies helpful, or the details of their implementation will make a difference to their effectiveness (e.g. sticking the sheets into Kevin's book).

In particular, it suggests that some pupils can use summative techniques formatively under certain conditions; Aaron is prepared to use levels and comments in feedback despite his rejection of levels alone, and Kevin seems to combine ego-focused and task-focused approaches to peer assessment. Thus, it is possible to find a 'more positive relationship' between the two paradigms (Black et al. 2003, 55, Harlen 2005), or rather that pupils can develop different strategies for developing a positive relationship.

Secondly, what forms of motivation and autonomy do different systems seem to foster? What is very clear from the two boys is the connection between their learning in religious education and the development of a personal view of the goals of the subject. Their appreciation of what to do next helped them to make sense of their own learning, and this motivated them because they could be task-focused (Gipps 1994). Moreover, their appreciation of the 'big picture' of the goals of the subject also contributed to this, because it related to them as

individuals (Munns and Woodward 2006). Motivation was linked to a sense of progress, but this was progress towards a wide understanding of the inter-relationship between knowledge, reflection and values; they did not separate the facts and the values. Self-reflexive educational autonomy (Black *et al.* 2006) developed in parallel with the development of a personal point of view and indeed with a sense of moral development (O'Grady 2003).

Thirdly, what does this show about pedagogy and assessment? The most important structural factor is the overlap between the subject-specific elements (notably attainment targets and attitudes) and the pedagogical requirements of formative assessment. The capacity to self-assess leads into the reflexive capacity and *vice versa*. 'Learning from religion' demands that the pupils reflect on their studies, and self-assessment demands the same. Both are meta-cognitive processes, which encourage the pupils to think for themselves, to be more independent and autonomous. This may mean that there is a difference between self-assessment in different subjects; self-assessment in religious education has a different nature to self-assessment in maths, which cannot offer these particular possibilities. The implications of this are further explored in the next chapter.

Fourthly, the main institutional factor is clearly that assessment for learning was a major part of the school's policy. It is strongly argued that a whole school approach is imperative (Black *et al.* 2003, James *et al.* 2007). This meant that pupils were likely to have encountered it elsewhere, and indeed were able to compare how different teachers introduced the same strategy, such as Kevin's remarks on model review sheets in science and religious education.

Nevertheless, it is striking just how endemic an ego-focused approach was even in an apparently exemplary school for assessment for learning (Looney and Wiliam 2005).

Fifthly, socially, pupils were clearly concerned with how they would appear to their peers; speaking as a teacher, I would say more than many other classes. The nature of the class's internal competitiveness and how well they co-operate as a group would also affect this. This was why the techniques were a 'culture shock' for Kevin. One can see how previous experiences of assessment have shaped their attitudes, such as their relative success or failure under more summative approaches, as well as psychological issues such as depression. At one level, it seems self-evident that a depressed pupil will be negative about themselves under a summative system, but what is less clear is why they would not be prepared to switch to a formative approach. In Jenni's case this was linked to her understanding of the teacher-pupil relationship, in which the role of the teacher was to play the game by setting banal tasks while the pupils tried to do as little as possible, but it may also be linked to a sense of the responsibility that autonomy brings – getting by is easier. Further, as the only girl, her example also points to the error in assuming that *all* girls have a positive attitude to their learning because girls currently generally perform better than boys; unfortunately, however, it is not possible to explore this question of gender further.

The assessment careers of these three pupils in religious education show us much about the impact of formative assessment on their motivation and autonomy. It seems that this notion of assessment careers is a powerful conceptual tool for diagnosing the ways that pupils respond to different

strategies. By focusing on the lives and experiences of the pupils, it provides a conceptual level on which generic assessment issues and subject-specific pedagogical ones can be inter-related. It thus also means that the effects on pupils of assessment strategies in different areas of the curriculum can be explored, e.g. why pupils related the use of formative approaches to revision for tests which they were never required to do, or the use of the same self-assessment strategies in different subjects.

More particularly, it helps teachers of religious education understand why some attempted strategies are unsuccessful with a class or with an individual, for instance the failure of the lesson plenaries. It shows that self-assessment and formative assessment more generally need to be differentiated, and thus not all strategies will be suitable for all pupils (see Laveault 2007). It also explains how individual pupils' reactions may be the result of long-term processes, which may help explain why they are difficult to change, or the result of what is happening elsewhere in their school life – secondary school pupils will be experiencing assessments with many different teachers (Natriello 1987, 111). However, as Aaron shows, it also can be the point at which they can gain the learning autonomy they desire.

10. 6 Conclusion

This chapter contributes to the originality of the research in two ways. Firstly, it uses biographical approaches to describe the 'assessment careers' of three pupils, Aaron, Jenni and Kevin (Ecclestone and Pryor 2003). This means that their individual voices can be heard more authentically than in the previous

chapters, and previous research. However, it thereby shows how pupils may shift between different approaches to their intelligence in response to different assessment techniques. In particular, it reveals how pupils can display a mix of both 'mastery' and 'learned helplessness' (Dweck 1999, 5). Furthermore, it shows how their sense of progression in religious education develops in relation to this shift, especially when they can develop a wider view of the goals of subject that includes all aspects of learning, toward learning autonomy. These two changes have implications for motivation. The chapter however also shows how some pupils will resist their teacher's attempts to encourage them to become more responsible for their learning.

Secondly however, through this approach the two theoretical fields of assessment for learning and the policy and pedagogy of religious education are inter-related in a new way. Patterns begin to emerge of how particular pupils make sense of their learning in religious education through the particular techniques and through their particular place on the spectrum from ego-focused to task-focused learning. Further, this process helps add rigour to the research because it creates another perspective through which to consider the data (Altrichter *et al.* 1993). The next chapter will consider the wider implications of the findings from all three chapters of analysis.

Chapter Eleven

Reflexive self-assessment as assessment for learning

The general issue at the heart of this study, as described in chapter one, is the nature of self-assessment in religious education. This chapter sets out to describe it in the light of the research. First, this chapter will summarise the conclusions reached in relation to the research questions. Then it will describe *reflexive self-assessment*, as an approach to self-assessment that is appropriate for religious education. Next, reflexive self-assessment will be compared with other theories of self-assessment. In the following chapter, the implications of the theory for the pedagogy and policy in religious education are considered.

11.1 From the research questions to the general aim

The general aim of this research was broken down into two general questions and these were translated into two research questions, each with a series of sub-questions (section 5.2). The first general question was about the classroom factors that affect self-assessment. The research question was more specific:

1 How can we understand self-assessment in the context of pupils' own ideas about their own learning and the classroom setting?

(a) How do pupils' views on different self-assessment techniques vary?

(b) Do pupils become more motivated as their ability to self-assess develops?

(c) What else affects pupils' identities as learners?

These are answered in detail in section 8.5, and a wide range of classroom processes are covered: particular self-assessment techniques; psychological profiles; social dynamics of the class; pupil-teacher relationships; experiences in other lessons. It is unsurprising perhaps that my pupils' comments on their learning matched the issues in other classroom research.

The second general question was whether the current attainment targets were adequately aligned with all the learning outcomes in religious education. The research question based on this was:

2 How do pupils consider that they improve at religious education, and especially 'learning from religion', as a result of self-assessment?

(a) What do pupils consider to be progress in religious education?

(b) Do pupils think that they have improved at this?

The research question is specifically answered in section 9.5, which suggests that almost all pupils could identify their progress in all three learning outcomes, learning about religion, learning from religion, and, importantly, the development of their attitudes and values. Furthermore, almost all were positive about their development in these areas. It is particularly significant that they could both

articulate a view of the wider educational outcomes of the subject, and express some sense of their progress towards this.

The second general question will not be fully answered by the research question, because the way that a subject is constructed, and its attainment targets formulated, is a more pedagogical issue than pupils are likely to be able to comment on. However, it would be foolish to construct pedagogies or identify attainment targets without paying attention to pupils' own views of their learning.

The data has been used to answer the research questions, and these answers inform the general questions. At this point, rather than considering the general questions first and then the general aim, the general aim is directly addressed and considered in the light of the general questions.

11.2 Reflexive self-assessment

Of particular note was the way that the different types of learning were intertwined in the pupils' self-assessment. Pupils did not necessarily distinguish between the self-assessment of values and the self-assessment of knowledge or reflection. The three elements were harmonised and integrated in their reflections on their different studies; they combine both intellectual as well as ethical qualities. This is a striking conclusion, but what should this form of self-assessment be called, and what are the implications of it for other theories of self-assessment and learning?

The most suitable title is 'reflexive self-assessment'. The term 'reflexive' is taken

from interpretive approaches to religious education, notably Jackson (1997, 2000b, 2004b, see above section 3.4). Borrowing it from the social sciences, he uses it to describe three processes: reflecting on one's own way of life (edification); constructively criticising the lives of those studied; finally 'developing a running critique of the interpretive process – being methodologically self-critical' (Jackson 2004b, 88). In this sense, reflexive self-assessment basically describes how pupils assess their ability: to reflect on their own lives; to constructively criticise others; to investigate the lives of others.

In order to explain this more fully it is helpful to look behind Jackson's work to discussions of reflexivity in the social sciences. This is to show the context of Jackson's usage, and also to show that my usage is not unwarranted. There are five points to make about the adoption of the term to describe this process of self-assessment.

Firstly, it is used to describe the process of pupils' reflection on *all* aspects of their learning. At one stage, in the social sciences, reflexivity was portrayed simply as a method of combating reactivity in research, i.e. the problem that the research itself affects and therefore invalidates the findings (McCormick and James 1988, 191, quoted in Cohen, Manion and Morrison 2000, 141), and so by being reflexive the researcher can mitigate these dangers. This is too narrow a view. By contrast, Giddens argues that it is a feature of all social life, as 'nothing is more central to, and distinctive of, human life than the reflexive monitoring of behaviour' (Giddens 1993, 120). Reflexivity is therefore intended here to mean a more general self-judgement of one's actions and thoughts, in the light of one's learning. The link between self and ethnographic method is made by Reinharz,

who suggests that reflexivity is to do with the researcher's *self* rather than the researcher's *role,* and identified a range of different aspects of her self that were important in conducting fieldwork (Reinharz 1997). Similarly, Abu-Lughod mentions the powers of personality and self-conscious methodology in shaping any research (Abu-Lughod 1988, 10). It is not just 'how could I have done this better?', but also 'what does this study tell me about how I shape myself?'

Secondly, it follows that reflection on values is also a part of reflexivity. In highlighting the values of ethnographic perspectives in inter-cultural education, Nesbitt links it to a reflection on ethics (Nesbitt 2004, 6). Jackson does not *explicitly* refer to values or ethics in his description of reflexivity. However, it is used within the context of Jackson's interpretive principles, because he takes account of the contribution of religious education to the development of values (Jackson 2004b, 161-180, see also Jackson and Fujiwara 2007). Thus, the argument here goes beyond Jackson's explicit text, but is nevertheless broadly consistent with his pedagogy. There is a philosophical dispute about the relationship between facts and values, e.g. the 'naturalistic fallacy', which cannot be examined here, though it would undoubtedly illuminate the problem (e.g. Moore 1903, MacIntyre 1983).

The components of reflexive self-assessment can be shown diagrammatically.

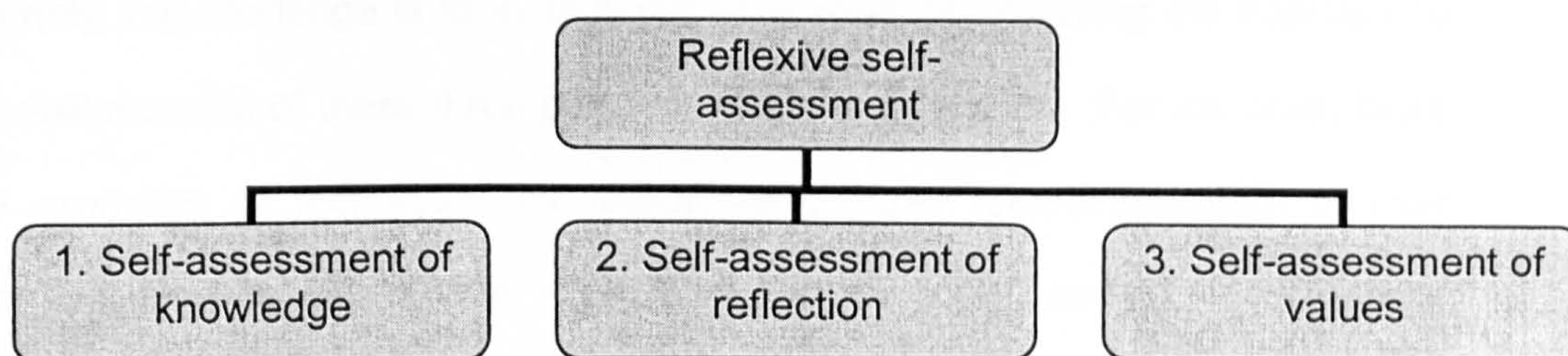


Figure 2: The components of reflexive self-assessment

Using the language of the Framework, the first element can broadly be described as self- assessment of learning about religion, and the second element can be described as self-assessment of learning from religion.

Thirdly, however reflexive self-assessment must be distinguished from 'learning from religion', or, in Jackson's terms, the two processes of edification and constructive criticism, as discussed in section 3.4. It is a higher order of reflection on these two processes, including such questions as: am I getting better at reflecting? Am I better at evaluating the views and values of others? Reflexive self-assessment should be viewed holistically. It is more than a collective label for the three kinds of self-assessment. The three elements neither coalesce into one nor are they completely separate, but are orchestrated into a whole that is distinct from its parts: 'a group of three pebbles could constitute a triangle even though none of the pebbles is triangular' (Honderich 1995, 371; holism entry). Without any one of them, pupils' learning would not be diminished merely because of its absence but also because of the consequential loss to the other two elements of that component.

Fourthly, the challenge is to identify effective ways of achieving the appropriate inter-relationship of these three parts for religious education. For instance, there are examples of techniques for encouraging pupils' self-assessment in their 'learning about religion' (AAIA North East Region 2003), and of how students' can self-assess their learning from religion (OFSTED 2003). Indeed, there are examples of pupils self-assessing their values in the Framework – even if they are not identified as such (QCA 2004, 6). All three types of self-assessment need to be more effectively harmonised and therefore all assessment and pedagogy in religious education should be re-conceptualised to achieve this. In particular, it means that assessment for learning needs to be fully developed for both intellect and values. This has implications for the use of levels within religious education both because of the immeasurability of values and because of the potential negative effects of using levels, as discussed in chapter eight.

Finally, this concept of self-assessment goes beyond existing conceptions of it in the assessment literature. This is not simply the incorporation of a theory from the generic research into religious education. It is the identification of a subject-specific strain of self-assessment as a result of grafting generic theories of assessment for learning onto the pedagogy of religious education. This has implications for both fields of literature. Therefore, following the order in which the two fields of literature were discussed at the start of the research, reflexive self-assessment is compared with other theories of self-assessment in the generic literature on assessment. Then, in the next chapter, its implications for the pedagogy of religious education are explored.

11.3 Comparing theories of self-assessment

The question that was raised at the end of the review of the literature on self-assessment (in section 2.6) was: what classroom factors affect self-assessment? Clearly, some aspects of this have been considered in the analysis in chapters seven and nine. The focus here is on the implications of this research for generic theories of self-assessment. Much of the research on self-assessment is from the mathematics and science classroom, but it may be that the particular dynamics of self-assessment in religious education can shed new light on it.

One of the criticisms of Black and Wiliam's original review article was that their conceptualisation of formative assessment was too simplistic, notably by French-speaking researchers (see section 2.4.5, and e.g. Perrenoud 1998, Allal and Mottier Lopez 2005). Drawing on Sadler's research (1989), they had identified three elements of which pupils need to be aware. These are: the desired goal; evidence about their present position; and some understanding of a way to close the gap between the two (Black and Wiliam 1998a, see Wiliam 1998). Black and Wiliam do not envisage self-assessment as simply the process of pupils' comparing their position with the desired goal – what would be tantamount to self-marking or self-scoring. While this may have some formative benefit (Maqsud and Pilai 1991), pupils also require an awareness of feed-forward. In the French research literature, however, a range of other aspects of this process are identified – e.g. self-determination (Laveault 2007, 213) the development of a structured concept of the self, and cognitive 'suppleness' (Campanale 2007, 198).

In addition, the critical issue for self-assessment in religious education is how this threefold process can be applied to the subject's various learning outcomes. I have discussed elsewhere some of the implications of this process in religious education (Fancourt 2005), notably the potential difficulties for pupils in developing a sense of this desired goal in their learning from religion and especially the development of attitudes and values: how do pupils develop an understanding of what tolerance or open-mindedness is? How could they envisage progressing towards them? The current research suggests that pupils can make sense of their progress in both these areas, and the implications of this finding within the debate about self-assessment needs consideration.

11.3.1 Understanding the goals

Firstly, what does this suggest about how pupils understand the desired goal? Learning from religion as a learning outcome is a way of allowing the pupils to connect with and make judgements upon the material that they have studied. Thus pupils begin to realise that they are not being judged on *what* they say, but on *how* they say it. The fact that there is no right answer can be seen as liberating, but it also begs the question as to what is required: pupils are expected to grasp that there are better and worse kinds of answers, better and worse ways of writing or saying what they think. For example, pupils could identify how to express their views more effectively,

> Dave: I have learnt how to argue both sides of it, of the whole aspect of the subject. [A1]

This ability is of itself not unique to religious education; it would also be important in English or history, for instance. Pupils are able therefore to understand what Sadler calls 'fuzzy' criteria in making qualitative judgements about their work (Sadler 1989, 135).

The more tantalising question is how pupils come to understand the desired attitudes and values. I described this as a major problem of 'plasticity' (Fancourt 2005, 118), using Wiliam's term (Wiliam 1998, 13). Many were able to make some sort of judgement about their progress towards tolerance and respect. But how did they come to develop their sense of this goal? Very clearly this is not arrived at through simply handing out criteria on a sheet, or even through the use of peer assessment, especially as there is no obvious type of 'work' that could exemplify it. The outcome is not necessarily 'in' the work, but arises holistically as a result of the content of the lessons (e.g. the Holocaust), the tasks set (e.g. to write an essay on whether British schools, like France, should ban the hijab)), and the activities (e.g. discussion), which all combine to infuse the pupil's attitudes to others.

Are there some external criteria by which to judge this progress? There are striking examples of complex quantitative research methods that attempt to identify changes to pupils' behaviour and values as a result of their learning (e.g. Narvaez *et al.* 2004, Tirri and Nokelainen 2007). In religious education in England and Wales, Smith and Kay (2000) have sought to show how pupils' attitudes and values are affected by different pedagogical approaches. However, the explicit use of quantitative methods in the classroom raises the spectre of pupils treating it in an ego-focused way, by treating the results as an indicator of

ability, and by reacting competitively. In the long term, more sophisticated methods of analysis may be possible, akin to the materials produced by the Center for Ethical Learning at Notre Dame University (e.g. Endicott 2001, Narvaez 2005); but they would of course have to be used in a formative way.

As we have seen, pupils understood these attributes and values in different ways. For instance, some pupils regarded tolerance as being close to sympathy for individuals:

> Mike: I'd say I have become more tolerant to the Jewish people who were affected by it. [B2]

For others it was connected to openness to others,

> Kevin: yeah - um, but doing discussions has made me think that ...it's about peoples' (*accounts?), as well, so I'm less intolerant. [D1]

One pupil focused on the 'condition of objection', in the sense that it was in response to a sense of dislike (Afdal 2007, 92):

> Alan: ...if you are saying 'tolerate them', it means you don't like them. [C3]

This might suggest that they have not understood or only partially understood the criteria. While however there needs to be clarity of language to describe attitudes and values, it is hardly surprising if the pupils' understanding of these

concepts is flexible, as is their sense of progress towards them. Mike's description of the limits to his tolerance reveals a concern to be precise about what tolerance is and how it has developed. He was not prepared to make a general claim that he was more tolerant, but delimits it to a particular group. This can be described as a 'thin' view of tolerance, as opposed to 'thick' tolerance which is more universalistic (Afdal 2006, 91). Alan highlights the condition of objection, and thereby casts doubt on it as a value: it is right to dislike people, even if you do learn to put up with them? This recalls academic debates about the distinction between dislike and disapproval, for example racial tolerance - do you have to dislike other races to tolerate them? (e.g. Nicholson 1985, 160-161, in Afdal 2006, 92). Alan interprets the criterion in a particular way, even if he then questions the value of the value thus interpreted.

11.3.2 Sharing the criteria

This raises a significant issue about the relationship between the teacher and the pupil in understanding the criteria. Sadler considers that it is the teacher's role to find ways of passing on his or her professional understanding of the criteria, however fuzzy these may be, to the pupil. While the pupil may become an independent self-assessor of their work, this is on the basis of the shared and accepted criteria; the teacher would always have foreknowledge of them. This approach could be seen as particularly appropriate in science and mathematics, where the content is less open to individual interpretation (see Cowie 2005, 146). It also underlies attempts in religious education to re-phrase the current assessment criteria in 'pupil-speak' (e.g. RE Online and Culham Institute 2006); in this situation, pupils are applying the externally specified criteria to themselves.

As described earlier, Pillonel and Rouiller describe this as an 'internalised other-assessment' (2001, see section 2.6). Pupils adopt the teacher's role, working to predetermined criteria.

When however it comes to sharing an understanding of attitudes and values with the pupils, the teacher's role is altered. This is not just because the criteria are harder to share; indeed this simply would mean that the teacher would have to work harder finding ways to share them, as is the case for 'fuzzy' criteria. It is also because self-assessment does not replace or even complement teacher assessment; for these outcomes, the main form of assessment *is* self-assessment. This may be supported or stimulated by the teacher in a variety of ways, such as through the use of traffic-lighting or discussion, but the main agents of assessment would be the pupils themselves. The teacher can only stand back and encourage the pupil to perform it. The teacher's role is that of 'adjustment' not 'regulation'[1] (Laveault 2007). However, whereas with most learning the teachers role is to move from regulation to adjustment, from control of the learning to supervision, in this situation it can only ever be the 'adjustment' of the pupil's internal reflections.

[1] 'Adjustment' is my translation of 'réglage', as opposed to 'régulation'. According to *Harrap's Shorter Dictionary* (1982), it is used to describe the process of setting or 'regulating' clocks or scientific apparatus, i.e. ensuring that mechanisms that are self-regulating continue to do so effectively.

11.3.3 Current 'attainment'

Even if pupils have both understood the criteria and are prepared to accept them, there is still the issue of judging and planning their progress: in other words, how can pupils decide how they are progressing in all areas, including how tolerant or respectful they are? The general issue was raised by a pupil in the pilot study:

> Dan: The thing is, how do you know what you don't know? [PD]

In terms of values it is even harder. Kevin thought he was more tolerant than he used to be because he could recognise a change in his attitudes to others in the class; he had some sort of experience to draw upon. By contrast, Mike considered that he was only more tolerant towards Jews who had experienced the holocaust. He had met a holocaust survivor who had come into school, but had no other encounters with them, so it would be hard to know how this would be revealed.

What is being demanded of the pupils is to be honest with themselves in a form of moral reflection. This is not easy for anyone. Furthermore, it would not be easy if pupils' other experiences of self-assessment were located within a more ego-focused paradigm. For instance, most pupils were uncomfortable with the use of levels for self-assessment, and for peer assessment, and they suggested strategies for avoiding the negative consequences of comparison by choosing levels strategically.

Mike: I'd say a lot of people wouldn't say what they thought in case they thought, "Oh I think I've got a level 6 but I'm not going to say it because people are going to shout stuff at me." [B3]

Obviously, there is no suggestion that pupils should decide what *level* of tolerance or respect they have, but there is clearly an issue in encouraging pupils to be both honest and open in the classroom about their progress reflexively. If pupils are used to 'playing the system', then they may well continue to do so, or rather, be unable to appraise themselves fairly. Pupils would need to be encouraged to be honest in the more measurable areas of attainment if they were to progress in all of them.

11.3.4 Closing the gap

If pupils have identified the target, and then identified their current 'attainment' in relation to it, the final stage is to have 'some understanding of a way to close the gap between the two' (Wiliam 1989). For the kinds of intellectual skills described by the current targets, these can be delineated; the problem is incorporating the third element: could pupils understand what they have to do to be more tolerant? Could they understand what they would have to do or show in the class to demonstrate it?

This could be seen as impossible, because the kind of progress envisaged may not be demonstrable in obvious ways. However, this is not a form of self-reflective behaviourism in which pupils could identify examples of more tolerant behaviour and then aim to attain them. Firstly, they would be almost impossible

to describe. Secondly, this would distort the pedagogy of the subject into a purely instrumental values education, and away from the reflexive inter-relationship of intellectual and ethical goals. Thirdly, at a practical level, if the kind of tolerance required was tolerance towards people from a variety of religious traditions, then those many communities and schools that lack religious diversity would not be able to provide the opportunities required for demonstrable action; for instance, there are no more than a handful of Muslims in Lord Williams's school to whom non-Muslim pupils could 'demonstrate' their tolerance.

It is also inappropriate because what is at the heart of the question is whether and how the pupils come to share these values. Kevin, for instance, considered that he had become less intolerant because he now valued other people's ideas, but would he have to recognise that he should aim to value them *more*? If so, how does one do that? David considered that he was tolerant, but was more strongly committed to the view that hijab was not necessary as a result of his studies. Arguably, he was being *more* tolerant because he had to tolerate something to which he was even more opposed. What is at stake is the volitional adoption of these shared values, but they do not translate into clear targets.

Nevertheless, it is possible for pupils to identify personal goals in this respect. In response to a question about 'learning from religion', one pupil described it as follows:

> Helen: you see, I've lived in many places before I came here, so I would respect a lot of religions, and I came with that. What I need to improve on

is whether I respect other people's opinions, and the way that if they get upset when I say my opinion, I will understand. [A2]

Helen identifies a complex 'target' for progression. It is not about denying her right to express her opinion in the face of diversity, but rather about anticipating the effects of disagreement. Nevertheless, this is not an argument to be allowed to say what she wants regardless of whether it offends others; her sensitivity belies it. Her target is not externally demonstrable, but is about the refinement of a disposition, to be enacted when the situation arises.

11.4 Reflexive self-assessment and the use of levels

One of the vexed issues in research on assessment is the inter-relationship between summative and formative assessment. Black and Wiliam had originally emphasised the negative effects of summative processes (Black and Wiliam 1998a), drawing on Butler's research (Butler 1988). This was criticised at the time (see section 2.4.4); Biggs argued for 'a broader context, embracing a multidimensional view of the instructional process' (Biggs 1998, 109). Black and Wiliam later suggested that teachers could develop the formative use of summative assessment, partly as a result of working with practising teachers who were constrained by the requirements of national assessment policy (Black *et al.* 2003).

In trying to reconcile this apparent tension, Harlen represented it as a 'dimension' (see section 2.4.4), though spectrum might be more appropriate. Harlen assumes that assessment activities only possess one function; however this

research has shown how pupils' understandings of the function of the activities varied, as Laveault suggests (Laveault 2007, 233). In particular, their assessment careers coloured their reactions; for instance the pupils' experience of working towards their Key Stage 3 SATs tests during the year coloured their reactions to assessment in religious education. At an individual level, Jenni reacted to any use of levels in an ego-focused way, whereas Aaron was able to distinguish between formative and summative uses.

However, the important point in developing a reflexive approach to self-assessment is combining the use of formal formative and informal summative purposes with the self-assessment of values. The problem here is that Harlen's dimension assumes that all learning outcomes can be assessed either way. While she has concerns with the *efficacy* of the different practices, i.e. that summative assessment may not provide the detailed information needed for feedback, or that formative assessment may not be reliable, there is no question that the criteria themselves are unchanged. This is linked to a tendency to envisage assessment as teacher-led (Harlen 2006, 117).

There is therefore a risk that reflexive self-assessment is impeded in two ways. First, too great a use of both formal and informal summative techniques will not only fail to provide pupils with feedback, but will also prevent the self-assessment of values. Secondly, even the use of formal formative techniques, e.g. teacher feedback, could also impede the self-assessment of values because these processes are teacher-focused.

11.5 Learning autonomy and its limits

This is linked to the question of autonomy. The relationship between self-assessment and autonomy is described by Pillonel and Rouiller:

> To develop in the learner a genuine actual ability to self-assess...is to dare to affirm their identity without ignoring their otherness. To risk letting the educated take the initiative should allow their progression towards autonomy. To accord to the learner genuine autonomy in their educational plan is to recognise them as a thinking subject; an indispensable recognition of identity so that they can develop themselves (Pillonel and Rouiller 2001, 5)[1].

It would therefore fall on the pupils to picture their own development in the light of their understanding of what it means to develop these attitudes and values. This is a significant reconstruction of the 'teaching contracts' (Perrenoud 1991, 92). It is impossible for the teacher to determine this feedback process; instead it is for the pupil to manage it with the teacher's support. Jenni for instance refused to take part in this process, and there is no way that I could have carried it out instead. If she does not do it, no-one does. This process will demand a high degree of active student involvement. Furthermore, it will demand creative and

[1] Translation NF: Développer chez l'apprenant une véritable compétence auto-évaluative en actes...c'est oser affirmer son identité sans occulter l'altérité. Se risquer à laisser le formé prendre des initiatives doit permettre son cheminement vers l'autonomie. Accorder à l'apprenant une véritable autonomie dans son projet scolaire c'est le reconnaître comme sujet pensant ; reconnaissance identitaire indispensable pour qu'il puisse se former.

subtle ways for teachers and students to negotiate assessment in particular in constructing shared understandings about what is expected (Allal and Mottier Lopez 2005, 245). The pupil's individual view of these outcomes may differ from the teacher's view. My personal view of what tolerance is differs from Mike's view and Alan's view, but their views are certainly within the ambit of accepted meanings for the word.

In the recent Learning how to Learn research, the importance of 'learning autonomy' has recently been highlighted (Black *et al.* 2006), as an overarching aim of the process of learning how to learn. The researchers highlight the fact pupils would have 'to exercise a degree of autonomy from the teacher as the assessor and judge of quality' (Black *et al.* 2006, 128); moreover, they argue that this is not,

> ...a single unitary concept, but rather that the term best represents a variety of overlapping and effective learning practices. In particular, it implies that the learner can not only give meaning to the learning, but that she can also create new learning tools. (Black *et al.* 2006, 129)

In terms of religious education, one of the major learning practices would be the inter-relationship between learning autonomy and the pupils' moral and religious autonomy; this is discussed below in relation to the pedagogy of religious education, below.

This however is not to argue that pupils can learn whatever they want from this process. I would be concerned if Mike was more intolerant as a result of my lessons, and indeed if he thought that he was more intolerant, or that this was a

good thing. The *shared* nature of the understandings is critical. There is a tension here between recognising that pupils are entitled to draw their own conclusions in learning from religion, but being clear about the values that underpin this process - what Jackson calls the 'non-negotiable' procedural values (Jackson 2004, 166).

The teacher's role in developing this form of self-assessment is in opening out the terrain, so that pupils can refine their own judgements. It is not a choice between autonomy and heteronomy, but rather tipping the balance in favour of autonomy. Ultimately no teacher can control what the pupils learn, and this will be particularly true in a subject that demands of pupils that they develop their own views and opinions. Nevertheless, certain outcomes are more possible than others, and form the educationally intended goals. The recognition of pupil autonomy in this sense is not a return to child-centred education, but rather the development of a more collaborative approach to the wider implications of learning. Pillonel and Rouiller's demand for 'genuine autonomy' is too strongly worded; it has the potential risks of a Sartrean demand for freedom from the coercion of others through 'bad faith' (Honderich 1995, 69). Autonomy should be situated, because 'individual agents cannot recognise or identify the contents of their minds without some sort of conceptual equipment with which to do so' (McKinnon 2003, 237), and moreover this conceptual equipment is not private, but arises out of shared understandings and 'forms of life', to use Wittgenstein's phrase (Wittgenstein 1958).

Religious education, like any curriculum subject, is what MacIntyre, terms a 'practice', as discussed above in section 4.1.3 in relation to research and teaching. This is a,

> ...coherent and socially established co-operative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially derivative of, that form of activity (MacIntyre 1981, 187).

A participant should accept of the standards of excellence for a practice, but at the same time 'the standards are not immune from criticism' (MacIntyre 1981, 190); they are open to question and re-evaluation. This operates at a number of levels; the pupils will themselves dispute the value and meaning of the virtues as will individual teachers and indeed academics. Nevertheless, the subject's rationale is founded upon its shared standards.

Pupils in my research clearly saw a connection between classroom discussion and a wider sense of 'listening' to the views of others, for example,

> Kevin:...doing discussions has made me think that it isn't, it's about peoples (*accounts), as well, so I'm less intolerant. [D1]

Moreover, the Learning how to Learn project also found that collaborative learning was one of the most effective ways of enabling individual learning autonomy, through reasoning and the questioning of ideas (Black *et al.* 2006, 127, James *et al.* 2007).

To suggest that pupils learn how to develop both explanations of behaviour and justifications of beliefs in a communal context is not to argue that religious education will achieve these instrumental goals in a straightforwardly behavioural way. Pupils will not emerge from the religious education classroom as saints: the subject will not *make* pupils good. It can however provide the arena for the development of a knowing reflection on the appropriate intellectual and ethical excellences.

It is certainly true that, in this context, other aspects of the classroom situation may be relevant as part of this wider negotiation. Pupils' views of their progress towards these excellences may be at odds with other aspects of the pupil's behaviour in the classroom: the pupil who indulges in prejudiced name-calling may think that she or he is very tolerant of different beliefs to her or his own. Consider the earlier example in context:

> Kevin: I admit, I admit I could be, a bit, er -
>
> Annie: pig-headed?
>
> Kevin: yeah – um, but doing discussions has made me think that it isn't, it's about peoples accounts, as well, so I'm less intolerant. [D1]

It is Annie who provides the critical judgment of Kevin's character - which he accepts as valid. Annie's judgement however is not made in a vacuum. It arises out of a situation – in this case the interview - that allows such reflection. What is also striking is that the social dynamics of the classroom can militate against assessment for learning (e.g. Cowie 2005), in this case, it aids a wider reflection on learning.

11.6 The wider context of reflexive self-assessment

This research suggests that as far as religious education is concerned, attitudes and values should be intermingled with intellectual skills in reflexive self-assessment. Laveault suggests that one should distinguish the processes of 'adjustment' that relate to the task from those what relate to the self, but he emphasises that those that relate to the self are more authentic, 'that is to say, at the level of the pupil's system of beliefs and values'[1] (Laveault 2007, 234). This is consistent with the development of attitudes and values as a long-term educational goal that will bear fruit in the future lives of the pupils as citizens in a pluralistic democracy.

The development of autonomy through self-assessment can also be linked to the issue of student engagement. As mentioned earlier (see section 2.4), the link between self-assessment and student engagement has been conceptualised at two levels by Munns and Woodward as: 'e'ngagement, which is the combination of cognitive, affective and behavioural factors at a routine level; and 'E'ngagement, which is that 'longer and more enduring relationship with schooling and education...the belief that "school is for me"' (Munns and Woodward 2006, 197). They argue that they are linked; the first is embedded in the second. Thus it is not simply about knowledge and ability, but also about control of the learning, about being valued as a learner and about having a voice, in an environment with 'students and teachers playing reciprocal meaningful

[1] Translation NF: c'est à dire au niveau du système des valeurs et des croyances de l'élève.

roles' (Munns and Woodward 2006, 210, see also Deakin Crick *et al.* 2007, 267, Marquez-Zenkov *et al.* 2007).

From the perspective of religious education, self-assessment can be understood as the development of an autonomous independent person, who can reflect on all aspects of learning in religious education. The various other elements both of assessment for learning and particularly self-assessment will still be important: questioning, teacher feedback, sharing criteria, pupils' self-assessment of their learning about religion, pupils' self-assessment of their learning from religion, and the use of summative assessment. However, this more unquantifiable potential would need explicit recognition too, and its inter-relationship with other aspects of the process developed.

Importantly, though, this reflection is not separate from the rest of pupils' learning in the subject. Their ability to decide if they tolerate, or respect, members of various religious traditions is bound up in their learning about, and learning from these traditions. Content-less moral education cannot replace religious education, because moral and religious judgements are embedded in context. The pupils' ability to make these judgements was as a result of studying the holocaust or *hijab*. The moral judgement is part of the factual learning; pupils do not learn the facts first and then decide what they think, rather they react to the material immediately in a multiplicity of ways. The pupils are challenged *both* intellectually *and* ethically.

11.7 Conclusion

This chapter has described how the data leads to a distinctively original model of self-assessment for religious education: reflexive self-assessment. The research set out in chapters seven, eight and nine was summarised to present an overview of answers to the research questions. This was then drawn on to address the general issue and general questions, on the nature of self-assessment in religious education. Reflexive self-assessment was defined as pupils' ability to reflect on a particular combination of intellectual qualities (i.e. the ability to learn from religion and learn about religion) and ethical ones, in terms of their attitudes and values. The background to the meaning of the word 'reflexive' in the social sciences was explained, and its background in the pedagogy of religious education (Jackson 1997, 2000b, 2004b), highlighting its intellectual and ethical dimensions.

Moreover, it is original because it goes beyond current conceptions of self-assessment in generic research. The difference between this and other theories of self-assessment were considered, particularly those of Sadler (1989) and Black and Wiliam (1998a), because these underpin English research and practice. The problems of understanding the goals, sharing the criteria, judging current attainment and closing the gap were explored. Then the inter-relationship between reflexive self-assessment and the possibility of using summative assessment formatively was considered; in particular the tension between the teachers' intentions in using summative assessment and the pupils' reactions to it. Harlen's research was questioned in this respect, for assuming a single function to assessment processes (Harlen 2007).

The discussion then turned to the ideas of learning autonomy, highlighting the privileged role of the pupils in reflexive self-assessment, and the implications of this for the teacher-pupil relationship, suggesting that this was not a simple process of handing power to pupils, since teachers never had control of this process in the first place, but that there were limits to what one should anticipate. This drew on Laveault's theories of 'adjustment' as opposed to 'regulation' of learning (Laveault 2007). Finally, reflexive self-assessment was placed within the wider issues of autonomy and engagement in the long-term learning goals of religious education. What however have not been considered in detail are the implications of reflexive self-assessment for religious education. The term is taken from interpretive pedagogy, but is it truly consonant with that approach?

Chapter Twelve

Reflexive self-assessment in religious education

If reflexive self-assessment is the appropriate model for self-assessment in religious education, then it should be aligned with pedagogy and policy. What would need to be in place for reflexive self-assessment to occur and develop? There are two parts to this discussion: firstly, the interrelationship between reflexive self-assessment and interpretive pedagogy is considered, especially in relation to recent research; secondly a review of the two attainment targets in current policy, which reviews the implications for a new Framework.

12.1 Reflexive self-assessment and pedagogy of religious education

My research suggests that there are three inter-related dimensions to self-assessment in religious education. At the level of learning about religion, teachers will need to share the factual criteria clearly, and both pupil and teacher assessment will be vital. At the level of learning from religion, teachers will need to share the 'fuzzy' criteria of edification, constructive criticism and expression; in this both pupil and teacher assessment will play a part, even if, obviously, the judgements themselves are personal. At the level of values, the criteria are even fuzzier and more open for pupils to interpret for themselves, and self-assessment will be central as teacher assessment is impossible. Self-assessment therefore needs to seen as a spectrum of practices from the teacher-led sharing of professional knowledge through to more autonomous judgements about all

learning, as the teacher's role moves from regulation to adjustment (Laveault 2007). The teacher is less the imparter of knowledge and more the facilitator of pupils' own learning.

12.1.1 Intellect and values

Reflexive self-assessment would obviously include pupils' ability to reflect on the development of their attitudes and values together with their intellectual progress. It therefore supports pedagogical approaches that acknowledge the development of attitudes and values alongside intellectual skills. For example,

> Penny: yeah, I think it's good to know you/ about different religion's views and things. Because people can be very like...narrow-minded against other religions, maybe because they have a slightly different belief to them - and I think that they are quite, it can sometimes turn into like maybe racism. [A3]

Thus Penny links knowledge, i.e. learning about religion, to the development of open-mindedness and anti-racism. Other pupils linked the development of values to their 'learning from religion', though as discussed above (section 11.3.1) this was sometimes in complicated tension with the expression of their own views, and their understanding of the anticipated attitudes:

> David: you understand their point, but it could either make me more tolerant or it could make you more against it, it's just completely, it's up to you isn't it? It's up to you. It's your opinion. It helps your opinion no matter what. [A3]

This therefore counters pedagogical approaches that explicitly reject attitudes and values as part of the goal of the subject, such as Wright, who contrasts an intrinsic value for religious education with an instrumental one. Religious education's intrinsic value is its 'ability to enable society to learn to ask fundamental questions about the nature of reality in an informed and intelligent manner' (Wright 2004, 216). An instrumental value is 'its extrinsic utility as a buttress to one or other modern or post-modern way of ordering society', i.e. its contribution to attitudes and values such as tolerance or respect (Wright 2004, 216). Jackson has contested this claim (Jackson 2004, 75-86), and O'Grady, from his research on applying an interpretive approach, argues that this dichotomy is too sharply drawn, finding 'pupils to be perfectly capable of combining regard for others with sharp views of their own' (O'Grady 2006b, 321, see also Gardner 1992).

My findings support the claims of Jackson and O'Grady. Many pupils were clearly able to understand the links between content, intellectual skills and attitudes and values, and reflect on how their progress in them was interconnected. More specifically, however pedagogy in religious education should set out its values explicitly: they should not be hidden, or assumed. *All* the learning outcomes should be shared. This makes explicit what is implied in Jackson's work, that pupils should reflect on all aspects of the interpretive process, including the values that underpin it.

Nevertheless, Wright is correct to criticise an imbalance in the relationship between the intellectual and ethical dimensions; to treat religious education as essentially values education would indeed undermine its educational justification.

This can be seen in the position put forward by Erricker and Erricker, who argue for 'non-epistemological education' (Erricker and Erricker 2000, 73) in which knowledge and intellectual skills play no part, but the important skills are those of 'emotional literacy', such as reflection on one's emotions, self-knowledge, reflection on the emotions of others, empathy, and recognition of difference (Erricker and Erricker 2000, 151). This is also too extreme. It is in the engagement that learning occurs, through intellectual and moral evaluation with different religious and non-religious points of view. Thus the argument for a *reflexive* approach to self-assessment is precisely to avoid an approach that eliminates or prioritises one element at the expense of others.

12.1.2 Interpretive approaches and action research

What is striking about this are other similarities with O'Grady's research. He investigated pupil motivation using an interpretive pedagogy in a secondary school classroom (see above section 5.1, and O'Grady 2003, 2005, 2006a and 2006b). Through action research, he studied the effect on student motivation of enabling the pupils 'to act as collaborative researchers with their teacher' (O'Grady 2003, 214). There are two observations to make before considering four points of similarity.

The first observation is that, although the link between assessment and motivation has been a recurrent theme in the current research, as it has underpinned much of the more generic research on self-assessment (e.g. Munns and Woodward 2006), O'Grady does not consider assessment at all. This is arguably an omission, since he would, as a teacher, have been making some

sort of judgements or indeed encouraging the pupils to make them. However, it is striking that a different perspective on motivation produces similar findings.

The second observation is that similarities may emerge because of similarities in pedagogy and research method. As mentioned earlier (section 5.4.1), my pedagogy was also broadly interpretive and dialogical. It is therefore likely that parallels would exist between my research and O'Grady's research. Moreover, we both used ethnographic practitioner research methodologies, although O'Grady's is more clearly *action* research, in that it had iterative cycles (see section 5.1). Nevertheless, it is striking that they do not disagree. Indeed, the fact that research in different schools into different aspects of religious education *does* coincide allows us to start to be more confident about the validity of both sets of data, as they start to suggest more generalisable conclusions (Maxwell 1992, 293).

Four of O'Grady's findings are pertinent. First, he found that pupils were keen to be involved in topic and lesson planning, describing it as being 'in control' (O'Grady 2003, 220). Thus,

> RE should develop pedagogies of active enquiry, enabling pupils to develop qualities of responsibility and participation. Methodologically, a balance should be sought between interaction with the teacher and interaction with peers. (O'Grady 2006b, 341)

This parallels my argument that reflexive self-assessment is part of the wider process of enabling pupils to become autonomous learners; it also echoes other

demands for more pupil involvement in lesson planning (e.g. Munns and Woodward 2006).

Secondly, his pupils described how they increased their self-understanding through both the content and the activities in the lessons; therefore,

> RE should include opportunities for pupils to evaluate the curriculum for its relevance to their existential or ethical concerns or to matters of personal significance to them. Teachers should build pupils' evaluations of the curriculum into future planning. (O'Grady 2006b, 241)

This is broadly consistent with my argument that self-assessment should include both intellectual and ethical considerations. O'Grady considers that teachers need to respond to the entirety of the pupils' educational experience and of the educational goals for the subject. His view is also consistent with the notion of 'adjustment' of pupils own learning.

Thirdly, his pupils were able to envisage a trajectory for their learning. They could describe changes as a result of their earlier learning: "as opposed to Year seven I have learned to question religion as a whole" (O'Grady 2006a, 321). But they could also anticipate the effects of further learning: "At the moment I think my ideas will change about life after death" (O'Grady 2006a, 321). These pupils had an understanding of their learning career in religious education, which they could describe and evaluate. They could do this about work for formal examinations: "as I go through GCSE work I will probably see the world as a whole picture" (O'Grady 2006a, 322). Strikingly, they did not simply see the

course as simply leading to a grade, but as a genuinely educational process with its own intrinsic value. This echoes the pupils in my research who could connect their learning with wider issues in their development, e.g. Helen's view of her change from year 8 to Year 9.

Lastly, he proposes that:

> The role of the RE teacher should be to collaborate with his or her pupils in the search for identity. The responsibility of the RE teacher is to model the criticism, creativity and responsiveness of the mature enquirer, and to be alert to the pupils' developmental needs as they change. (O'Grady 2006b, 342)

My research would support this proposal, but assessment processes need to reflect this collaboration as well. There is no eight-level scale in 'the search for identity'. However, if "school is for me", as Munns and Woodward demanded (2006, 197), then it needs to be for "ME"; it needs at some level to explore who pupils understand themselves to be. This is autonomy, approached from a different direction. Black *et al.* argued that learning autonomy is learnt through collaboration (Black *et al.* 2006), and O'Grady suggests that the teacher joins the pupils' *own* search for identity, like a developed form of Davis's 'hermeneutic' assessment (1997). As I have argued, reflexive self-assessment is not a delegated form of teacher assessment, but is rather a form of assessment of the self that the teacher 'adjusts' (Laveault 2007), using his or her professional skills.

Both O'Grady's and my research show that pupils can indeed combine the various dimensions of learning in religious education. They can both hold strong views and reflect on the implications of this for the development of attitudes and values. They can develop a wider vision of the educational value of the subject even within the constraints of studying for examinations, and indeed this is an important factor in their motivation. Moreover, his research is consistent with my argument that reflexive self-assessment is vital for religious education, and that to achieve this requires a restructuring of the relationship between teacher and pupil: a re-negotiation of the learning contract as Perrenoud demands (1998).

12.2 The attainment targets revisited

What are the implications of this research for the current attainment targets? It would be tempting at this point to produce new revised attainment targets in the light of the research that governmental bodies and therefore teachers should adopt. This however could be seen as naïve and hypocritical.

It could be naïve because, as has been explained in chapter two, there are three contexts to the development of policy. New revised targets would be simply a further voice in the context of influence, which might affect the process of text production, but probably have less impact elsewhere, for instance in the political sphere.

It could be hypocritical because this current research is itself the result of a 'writerly' approach to the texts, to use Bowe *et al.*'s terminology from Barthes (see the introduction to chapter three, and Barthes 1975, Bowe *et al.* 1992). This

means that I have been able to interpret and therefore implement the texts in a creative way, through the research. The existing statutory and other official documentation is open enough to allow this research to be carried out: there is enough ambiguity between the Framework, the Guidelines, Oxfordshire agreed syllabus and OFSTED reports, and also non-statutory official texts encouraging assessment for learning (e.g. DfES 2003), to allow me, as a teacher, to treat it as a 'menu' not a 'mandate' (Hargreaves 1997).

If the documentation was more 'readerly', with an unambiguous 'author-itative' meaning, especially in terms of assessment, then this would be harder to achieve; my job would be the delivery of what was in the text, including the use of a prescribed model of assessment - and I would be Kafka's beetle (Parkison 2008). In the light of these concerns, what therefore is more appropriate is a review of the nature and shape of the document, rather than a re-drafting of its details.

12.2.1 The context of text production

Firstly, any document that is more harmonious in approach would need to be a more writerly – a 'menu' (Hargreaves 1997). There are two aspects to this. Any text is open to readerly or writerly interpretations, but some more so than others. Texts can be intentionally readerly, but interpreted in either way; Barthes himself showed how an apparently readerly text was open to multiple writerly interpretations (Barthes 1970). By contrast, they can be intentionally writerly, but interpreted in either way; it might though be more difficult to interpret a writerly text in a readerly way. To recognise that a text is writerly is also to recognise that

it is itself the product of the other texts, in this case both texts from the context of influences, e.g. political, academic, faith communities, and from other formal documents, e.g. the National Curriculum. In any event, it is important to recognise that the result of text production is rarely a simple unequivocal meaning 'but a multi-dimensional space in which a variety of writings, none of them original, blend and clash' (Barthes 1977, 146, Culler 1983, 78-90).

Further however, to approach the text as writerly is also to recognise that its interpretation and implementation will depend on the context of practice, a context that involves recognising the place of the pupils as well as that of teachers. Nevertheless, the meaning of the text will only be transmitted through the qualities and attributes of the teachers, as they edit it for their audience. This means recognising the autonomy of the practitioner. The issue of pupil autonomy has already been discussed in relation to self-assessment, but the issue of teachers' professional autonomy is also relevant here. A writerly text would have to take into account both pedagogical issues associated with religious education, the various assessment paradigms in operation, and open the way for independent creativity by teachers in selecting the meanings and translating them into the classroom. At a political level, this echoes Hargreaves' claim that policy-makers should replace the quest for a structural solution with a more cultural one, in the sense of cultures of practice; it would be about seeking to change classroom practices from within, and not from without (Hargreaves 1997, 350).

However, professional autonomy, like learning autonomy, is not limitless. Firstly, any text will have only a limited range of meanings. Secondly, the deregulation

and trust that this approach implies is matched by the requirements of knowledge and responsibility.

> It is because professionals face complex and unpredictable situations that they need a specialized body of knowledge; if they are to apply that knowledge, it is argued that they need the autonomy to make their own judgements. Given that they have that autonomy, it is essential that they act with responsibility – collectively they need to develop appropriate professional values (Furlong *et al.* 2000, 5).

The fact that teachers could interpret texts in a multiplicity of ways does not mean that any interpretation will do. Rather it is to recognise in the context of text production that the central features of the network of threads should be delineated, in order to allow a responsible reading of them, recalling Wintersgill's remark that such texts should be 'broad and flexible enough to allow for interpretation' (see section 3.1.1; Wintersgill 2007).

12.2.2 Problems with the Framework

Given this need for such an approach, what are the problems with the current documentation? The main problem, in the light of the current research, is the differential treatment of the development of attitudes and values in comparison with the two attainment targets, in particular, the separation of 'attitudes in religious education' from 'knowledge, skills and understanding' (QCA 2004, 13). The separation is predicated on a distinction between the cognitive and the affective, and is drawn from other parts of the National Curriculum; however, it is

precisely these attitudes that Clarke and Boston - the politician and civil servant in charge of education, respectively - considered that religious education should develop (QCA 2004, 3). This research has shown that pupils are able to identify how their attitudes have developed as a result of their learning.

As described earlier, many academic critics have argued that the two attainment targets should be combined (section 3.2, see for instance Jackson 2000b, Copley 2005, 126-128); very recently OFSTED have argued the same point (OFSTED 2007, 38). My data also support this. However, an even wider perspective is required; it is not simply that the current targets should be pedagogically combined, but that more widely all the dimensions of learning in the subject should be treated with parity.

This difference in status is particularly apparent when the attainment targets are described. The level descriptors of the two targets are the result of much critical thinking and drafting. They tend towards hyper-specification, which is another way of saying that they are very readerly in intention. Moreover, there are no level descriptors for the development of attitudes and values. At one level, this appears almost banal, since such development is notoriously hard to measure, however there are two effects of this omission. First, it means that such development is excluded from the readers' understanding of the object of assessment. Assessment applies to the cognitive attainment targets, which thereby have a higher status, and almost incidentally some attitudes and values are apparently involved. In reality they should be assessed too. Secondly, the omission reinforces a summative model of assessment, because it is inextricably linked to levels; the inference is that assessment is only possible with levels.

In short, the National Framework currently works within a paradigmatically limited framework for assessment in religious education. It points to a wider vision of pedagogy, but then does not develop as wide an understanding of assessment as this would require. It is true that it states that 'apart from their summative use, these level descriptors can be used in assessment for learning' (QCA 2004, 34); however this presumes that the primary purpose of assessment is summative, and the central method of assessment is the use of levels. This in particular means that the targets are intended for teachers' use; thus, 'in deciding on pupil's level...*teachers* should judge which description best fits the pupil's performance' (QCA 2004, 35 [emphasis added]). A similarly restricted understanding of assessment can be seen in the recent OFSTED survey; it was keen to encourage the use of levels for monitoring attainment and target setting (OFSTED 2007, 41).

This is not to say that the level descriptors should be abandoned altogether. Although they have potentially negative effects, pupils can, under certain conditions and given certain assessment careers, use levels formatively, as this research has suggested, notably Aaron. The levels, their uses and the consequences of these uses are not the same. Indeed, one can imagine a pupil using the levels formatively despite the fact that her teacher actually does not fully understand them or use them successfully with other pupils. They may also be useful to teachers for diagnostic and evaluative purposes, though that does not mean they always need to be shared with the pupils themselves. It is a question of responsible professionalism. The argument is not either/or but both/and; it is about the development of more nuanced menus that engender *more* responsible readings.

12.2.3 Redrafting the Framework

There are three implications for any new Framework. Firstly, re-phrasing the mandate is unlikely to establish the culture in which a menu can be used responsibly, to use Hargreaves' terms (Hargreaves 1997). In fact, a cause for complaint by OFSTED is not simply that teachers do not apply the levels, but that they do not use them to inform planning (OFSTED 2007, 22-23); their frustration therefore is not with the levels themselves, but rather with the alleged lack of a professionalism that can respond to the implications of the data. No amount of re-drafting will overcome this alleged failure to use the documentation properly.

Secondly, this also applies to other suggestions for a new mandate. Everington argued in 2000, that teachers felt 'directionless' and 'despair at the complexity and enormity of the task' because policy documents (specifically, the Model Syllabuses) contained three aims: understanding of religions, respect for difference, and personal, moral and spiritual development (Everington 2000, 193). She suggests that religious educators should come together to decide what the aim is, and then develop an appropriate pedagogy. This is certainly an argument for greater control of the curriculum by practitioners; however, this never goes beyond a mandated approach, even if it is more consensual. It would still be decided upon and then set out as what religious education should be. The problem is that teachers are likely to feel disenfranchised if whatever they are asked to deliver does not allow them professional autonomy. Even if an aim was 'agreed', there would still be those who did not accept it, or accept all of it, or find it does not apply. Further, any curriculum area is part of the wider experience of schooling, and therefore a number of stakeholders are implicated.

There are many voices, and types of voice, in the context of influence, not least faith communities.

Thirdly, it is not necessarily about eliminating ambiguity from the aims of the religious education, as Everington suggests (Everington 2000, 193). Instead, it is about recognising and encouraging the creativity of teachers in drawing on these different strands in their teaching. If the subject were reduced to one of the aims that she identifies, it would arguably be poorer for it. The issue is not the complexity and enormity of the task, but rather the nature of accountability. Hargreaves argues for a 'reconstruction of personal trust', such as the recognition of 'idiosyncratic excellence', to replace a misguided trust in overly-bureaucratic processes (Hargreaves 1997, 347). Teachers should not necessarily seek to *deliver* a simple curriculum; they could *engage* their pupils with a complex one.

12.3 Conclusion

In this chapter, the implications for the development of reflexive self-assessment in religious education have been explored more fully; thus, it develops the ideas of the previous chapter consistently. Its impact on the relationship between learning from and learning about religion, and their relationship to the development of values, was described. Jackson's approach was broadly confirmed in recognising the ways that religious education can contribute to the development of attitudes and values, against Wright's argument that the subject should not have an instrumental value. Furthermore, it confirmed an approach

that was rooted in religious material, against Erricker and Erricker's argument that this should be replaced (Erricker and Erricker 2000).

In particular, four similarities with O'Grady's research were identified: pupils should have more control of their learning; teachers should respond to pupils evaluations of their learning; pupils could describe a vision of their learning; teachers should work collaboratively with pupils in their search for identity. The similarities between on one hand this vision of pupil engagement and the pupil-teacher relationship and, on the other hand, the ideas of Munns and Woodward on motivation and of Laveault on the pupil-teacher relationship were emphasised.

The practical implications of this research for the current policy documents were then set out. They showed how there is a tension in the current documents between the desire to inculcate certain attitudes and values and the way that the attainment targets and the eight-level scale are formulated. A contributory factor is the unresolved status of the text as both mandate and menu – to be dutifully followed and to be interpreted freely. It is suggested that future versions should be explicitly writerly in style, allowing for a responsible professionalism to select as appropriate in response to classroom practice. This contributes to the practical significance of the research, as well as re-appraising the original use of policy scholarship in the analysis of policy documents in religious education. However, at this point, it is indispensable to review how successful the research has been.

Chapter Thirteen

Conclusion: reflections on the research

The previous chapter considered the opportunities for and challenges to reflexive self-assessment in religious education. This is a far cry from the research lessons, and it would be naïve to be overconfident about generalisations drawn from sixteen respondents. The research has grown out of my practice as a teacher, and so it is appropriate to reflect finally on this interrelationship. There are four areas to consider: the research itself; the nature of the project as both research and practice; its effect on my practice; what future developments might be. Finally, the quality of this investigation is defended.

13.1 Reflecting on the research process

It is typical in any research to review the process, and this in itself is a reflective activity. It is the researcher's own 'reflection on action' (Schön 1983). No researcher ever has research evidence *outside* the research for the success of the research; one can only say how one thinks that it did or did not follow the research design and what the consequences of any perceived omissions are. It is a subjective reflection against the epistemological norms of the research methodology.

How, therefore, with the benefit of hindsight, would I have changed the research itself? In general, the fact that I am broadly confident with the findings is a reflection of the fact that I am broadly confident with the research process. The first area of regret is that I did not identity the literature on pupils' experiences of

assessment, notably Ecclestone and Pryor's notion of assessment career, till after the research was under way (Ecclestone and Pryor 2003). It would have provided a useful starting point for the formulation of the research questions, and would have provided a rich area to explore in the interviews. At the time, I did not probe pupils on their references to other subjects, tests or exams. This could have provided a wider landscape within which to situate their perceptions of assessment in religious education.

Secondly, I am concerned with the triangulation of the data, notably the lack of any other data apart from the interviews. This is a function of two decisions. The first is the decision not to use any form of summative assessment in the research as evidence of progress, and the second is the decision not to continue with the questionnaires. While I am content with each decision individually, it does mean that there is no direct triangulation of the interview data. This problem could have been addressed in other ways, perhaps through asking an independent researcher to observe some of the lessons.

13.2 Reflecting on the practices of research and teaching

More broadly, how did this research and my teaching inter-relate? In terms of the epistemological basis of the research, virtue theory has much to offer as a research paradigm for practitioner research. This is because it can allow teacher researchers to balance the epistemic virtues with the ethical ones in a situated way for both the practice of teaching and the practice of research. It is neither simply about performing certain processes that arrive at knowledge (under a positivist or constructivist paradigm), nor about arriving at transformed social

relationships (under critical theory). To arrive at making a judgement about how pupils reflect on their learning meant making a balance between scepticism and credulity, between the needs of individuals and the needs of the class, and between self-effacement and intellectual arrogance, all at the same time. For example, in terms of the details of practitioner research, there is no reason to suppose that iterative cycles, a distinctive feature of *action* research (Elliot 1991), are essential. They may be useful, and it may be helpful to repeat research in other settings, but that is true more generally of all research and not just practitioner research. They are one technique for avoiding either scepticism or credulity, but are not an essential one.

Further, in terms of methodology, my research suggests that practitioner research has a serious and substantial contribution to make to the investigation of classroom practice. For the reasons outlined in chapter five, it interrogates the interplay between practice and theory. Afdal describes this as an 'abductive' approach (Afdal 2007b), also using virtue theory to identify the two 'practices' (MacIntyre 1985, 187). The practice of research informs the practice of teaching, but the practice of teaching also informs the practice of research. My research is about an attempt to introduce a form of self-assessment (drawn originally from Black and Wiliam 1998a) into religious education, but the practical consequences of this led to a reformulation of the theoretical understanding of self-assessment for religious education. With Afdal, I would maintain that neither has priority. The challenge for teacher researchers is not simply holding to this, but living it out professionally.

Moreover, Afdal has recently argued that it is in this engagement between the two practices that new understandings are created. He draws on Engeström's development of activity theory to argue that 'when the theories of research meet the theories of practice, the ground is prepared for valid theoretical expansion' (Afdal forthcoming, Engeström 2005, see also Vygostsky 1987). This research has shown how a new theoretical understanding of self-assessment arises out of the meeting of not simply theories of research and theories of practice, but moreover two different areas of research theory, assessment for learning and religious education. This new understanding has implications for both fields of research and for practice.

Even if the process is essentially sound, there were clearly aspects of the process that affected its quality (see section 6.3). One example, in terms of my theoretical understanding, was a lack of a full awareness of much of the research literature in French until the research was under way; this would have refined the questions considerably. Another example, of a research process, was the decision not to continue with the questionnaire data; this would have given a useful insight into the whole class's perspectives, and not just the interviewees. This was especially limiting given the decision not to carry out any summative testing in the research, because it meant that there was no triangulation of the pupils' perspectives, even if there was triangulation of the coding process by testing for inter-rater reliability. A final example, from the teaching, was my lack of determination in not seeing through the use of the rapporteur – this could have developed a more habitual process of pupil assessment. Nevertheless, despite these shortcomings, I am largely satisfied with the process. What the pupils say, and the depth with which they say it, is clear and unequivocal.

Further, the accumulation of different pieces of practitioner research allows one to compare and contrast them. My research, in a different school on a different issue, reached similar conclusions to O'Grady's research. This means that both results can be treated with more confidence, as being more generalisable, but of course it also leads one on to consider where and why they are different, and to interrogate the effects of similar methods and pedagogies. By placing the research in various relationships with other research, one can start to identify strong connections and new relationships.

13.3 Reflecting on teaching while researching

Clearly, this research did not happen in a vacuum. Even when I was carrying it out, I was teaching other classes, with whom I would implement or modify the approach. For instance, a Year 7 class that I was also teaching responded better to the rapporteur technique, the year group that the Learning how to Learn project suggest is more amenable (Learn how to Learn 2002). Since then I have taught several other Year 9 classes, over the same syllabus. The research has led to a wider range of reflexive questions to the students on their learning.

The ideas could also be extended over other classes, into GCSE and A level classes, where the tensions between the summative nature of formal examinations and the other dimensions of the subject are more pronounced. Thus with these classes, particularly GCSE classes, it arguably becomes more critical to draw attention to other aspects of learning than grades, to balance the 'hard demands of economic capability' with the 'soft equipment for the good life'

(Haakedal 2007, 49). One consequence was the drafting of a set of principles for assessment to balance these tensions (appendix 12).

In terms of the texture of lessons, the research has shown the inter-connectedness of different kinds of learning, and this impinges on the way that I introduce ideas. Before, because the values and attitudes were not in the assessment criteria, I tended not to make them explicit, even if I thought that they were important. That is no longer the case. This is in particular as a result of the interview process itself as much as a result of the analysis. Teachers do not usually have the opportunity to have long conversations with small groups of pupils about their learning. I was able to gain an insight into the way to phrase and frame questions in the classroom. Thus it was a master class for me in questioning techniques. Whereas before I aimed at identifying challenging intellectual questions rooted in the attainment targets, I now would consider raising a question about tolerance or empathy or open-mindedness, such as whether coursework on marriage in Islam has made pupils more tolerant. These are considerably more personal and richer questions, as O'Grady and Laveault would suggest (O'Grady 2006b, Laveault 2007), and in retrospect I can trace the development of my questioning, but beyond the form envisaged by Black and Wiliam (see Black and Wiliam 1998a, Davis 1997).

Unlike an academic researcher, there are implications in building up relationships with the pupils through the interviews. I know more about these particular pupils than I would otherwise have known. I have taught some of the class for GCSEs, and A levels. Our teacher-pupil relationship was inevitably affected by this. It is difficult to judge whether this was to the pupils' benefit or not. With some of

those individuals whom I did continue to teach, the content and nature of the research became part of our ongoing dialogue, both in terms of shared understandings about the subject and their learning. For others, it seemed to disappear from the horizon.

Moreover, there was a tension at times when particular pupils were being discussed by other teachers: how much could I put forward what I knew about them from the interviews? Clearly, on the one hand there was an expectation of confidentiality, but on the other, the information might help other teachers understand the particular pupil's attitudes to work and assessment. Further, how could I distinguish what I gathered while teaching them from what I gathered while interviewing? The first kind of teacher's knowledge would have been freely shared, and indeed it would almost have been seen as unprofessional not to share it, and the second was potentially confidential, but where did one end and the other begin? Clearly, if there had been a major revelation, then I would have passed it on, but these were more ordinary issues of school life. Much depended on the context – was another teacher struggling with the particular pupil, so that the research-based insight would have eased their relationship? It would be superficially easy to say that one should do whatever was in the best interests of the pupil, but that is hard to judge at the micro-level.

Is it possible to make sense of these dilemmas? I argued that practitioner research can be situated within virtue theory. This dilemma can be understood as the need to find the virtuous balance between the vice of secrecy and the vice of indiscretion. The problem is of course balancing this with other virtues, of professional compassion for the pupils and shared collegiality with other

teachers. In this context, what is important is to recognise that, whereas academic researchers leave the field, practitioner researchers do not, and this has implications for the practitioner researcher's professional life and relationships after the research has been completed.

13.4 Future developments

There are inevitably a number of different ways in which this research could be developed. The first and most obvious is in trialling a model of reflexive assessment to see how effective it is. The theory of reflexive self-assessment arises out of the research, in answer to the research questions, but was not itself the 'starting point' (Altrichter *et al.* 1993, 33). This trialling could be seen as the second phase in an action-research iterative cycle. This could either be me repeating the process, or other teachers, either in the same school or in other schools, similarly to the Westhill and FARE projects (Westhill College 1989, Copley *et al.* 1991, see also Torrance and Pryor 2001). Furthermore, this could focus on other possible influences, such as gender, or ethnicity, or the religious background of the pupils. Do girls find it easier to do than boys? Does it have greater impact on boys rather than girls? Are there cultural or religious factors that impede or facilitate pupils' abilities to reflect in this way? It would be interesting to see how it could be implemented in schools with pupils of a different ethnic and socio-economic background.

The interconnections between reflexive self-assessment and pedagogy could also be investigated further, which I am doing through the REDCo project. As mentioned above (section 10.6), this is a major European research study,

comprising of a series of inter-connected investigations across eight countries into the way that religious issues are handled in school (Jackson *et al.* 2007, particularly Weisse 2007). The British contribution to this is through a 'community of practice' using action research to apply and develop the interpretive approach (Jackson and O'Grady 2007, see also Wenger 1998). There are a number of individual studies looking at a range of issues including gender, representations of religion in the media, and teacher training. My particular contribution is on the implications of assessment for the interpretive approach. Some of these issues have been explored in this current research, though the main difference is that the current research has been concerned more narrowly with self-assessment. The REDCo research will allow for exploration of the implications of assessment paradigms within a particular pedagogy.

It would be wrong nevertheless to assume that self-assessment can only be manifested within interpretive pedagogies; if it extends across the curriculum, it extends across pedagogies. Grimmitt's original conception of self-evaluation was within a human development model (1987); it could also be applied within Wright's 'critical pedagogy' (2000). Research on the nature of self-assessment within these different approaches would also elucidate whether my findings are limited to an interpretive pedagogy, or to religious education more generally. One would expect that a different pedagogy would produce a differently nuanced form of self-assessment, because all assessment needs to be embedded in the epistemological and ethical presuppositions of the subject, otherwise, as has been shown, pedagogy and assessment are out of alignment.

Another thread that could be developed is the nature of judgements about values within pupils' assessment careers across the curriculum. There might be other subjects where this kind of approach could also be developed. In citizenship education, Pike, for instance, argues that 'in a liberal democracy it is likely that a degree of complexity will characterize its integration and assessment' (Pike 2007, 227). It would be unsurprising given the essentially social nature of the values demanded that such an approach to assessment would be possible in this area. He has also highlighted the role of moral (and spiritual) education in English and art (Pike 2002), and Hopwood has argued for recognising the place of pupils' perspectives of values in geography (Hopwood 2007). There are a number of other places in the curriculum where pupils could reflect in this way, if they had the opportunity.

This thread could however be wider still. Throughout all stages of education there are increasing demands for the development of ethical awareness among pupils and students in higher education, e.g. at IDEA-CETL[1]; it is particularly recognised in vocational studies, for instance in legal education (Economides 1998). But if those students have succeeded through an essentially valueless ego-focused approach to assessment, then not only would they display what Dweck described as learned helplessness (Dweck 1999), but they would also struggle to make wider ethical decisions about their learning.

Finally, the issue of spiritual development has not been discussed. The topic is nonetheless part of the debate about the role and purpose of religious education (e.g. Education Reform Act 1988, section 1(2a)). It was included in the Model

[1] http://www.idea.leeds.ac.uk/

Syllabuses, in the end of key stage descriptors, and is also mentioned in the Framework (QCA 2004, 14). It was omitted because the topic is an extremely complex one, and to do it full justice would have meant tackling a completely different field of research (e.g. Wright 1998, Copley 2000). Nevertheless, reflexive self-assessment potentially offers a way of developing it in religious education, and possibly more widely across the curriculum.

13.5 Contribution to educational research

To identify ways that research could be developed is to some extent to recognise its limitations; to conclude, therefore, the research's contribution to educational research is highlighted. As set out in the opening chapter, three aspects are considered: significance, originality and rigour. These are, ironically perhaps, taken from a current governmental 'assessment' of British research, the Research Assessment Exercise ('RAE') (RAE 2006). The conclusion of each chapter from two to twelve has explained that chapter's particular contribution.

The significance of this research lies firstly in its academic importance. Issues of assessment, attainment and motivation are central to much educational research, as set out in chapter two. This research has shown, in line with much research on assessment for learning, how a summative paradigm for assessment can in fact have a negative effect on pupils. It has also shown how pupils, such as Aaron, welcome different approaches, which allow them to create a new sense of their learning career. It is also significant because it tackles the issues of how teachers can assess the immeasurable, i.e. pupils' ability to reflect on their attitudes and values. It shows that pupils are able to develop their own

understandings of complex criteria, such as tolerance, and develop the ability to apply them to themselves. Further, it tackles a key pedagogical issue in religious education of how the knowledge and understanding of religions can be linked to both the ability to reflect, and the development of the appropriate attitudes and values for pluralistic democracies (see chapters three, eleven and twelve).

Its significance also lies in the fact that these are current practical problems. The debates about on the one hand, the nature and purpose of assessment in schools, and on the other, the nature and purpose of religious education are not simply theoretical. They go to the heart of classroom practice, and are the centre of heated discussions in the teaching world (as described in chapter three). This research suggests that the classroom assessment in religious education could be very different. The role of the teacher would be as a more autonomous professional, seeking to aid the pupils' personal intellectual development and the development of their attitudes and values. There would be less reliance on the eight-level scale as the main or only arbiter of progress, and policy documents would be more like menus and less like mandates. As a result, pupils could make more progress because they would be more independent learners, able to integrate different aspects of their learning together. Furthermore, self-assessment across the whole curriculum would not just be about the ability to succeed in tasks, but be more holistically nuanced towards their selves, giving pupils more opportunities to integrate their intellectual and ethical development (and perhaps their spiritual development as well).

The research is also original. This is firstly because it 'engages with new or complex problems' not 'straightforward' ones (RAE 2006). One aspect of this is

that it explores the inter-relationship between the generic research literature on self-assessment and the pedagogy of religious education; this of itself makes it unique. Another important aspect of its originality is the exploration of classroom assessment of attitudes and values. Lastly, and most importantly, it proposes a new way of describing this interrelationship in reflexive self-assessment. This has implications both for the pedagogy of religious education and for generic theories of self-assessment as assessment for learning.

Secondly, it is particularly distinctive because it tackles the problems in new ways. Most importantly, there is no evidence of other empirical research specifically into self-assessment, or assessment for learning, in religious education in England and Wales. More particularly, the literature review of the policy and pedagogy of religious education uses policy scholarship to analyse the debate (chapter three). Moreover, there are few examples of practitioner research or action research in religious education. Finally, the articulation of a more nuanced approach to practitioner research methodology through the use of virtue theory is especially new (chapter four).

The last aspect is rigour. Significant, original research is flawed if it is not rigorous. This research combines the demands of research with the demands of teaching, so that both areas of practice are sustained (chapter five). It is also rigorous in that, within a virtue theory paradigm, I use a range of research tools to justify ensure that my findings are justified, such as detailed transcription, careful coding, inter-rater reliability, and handling of reflexivity (see chapters five and six). In particular, it means that one can be confident that there was no negative influence by my being both teacher and researcher.

Overall, therefore I believe that this research strikes effectively and creatively at the heart of a much-discussed but hitherto uninvestigated dilemma. The notion of reflexive self-assessment potentially offers a novel solution for both theoretical and practical issues facing religious education and assessment. These could be developed in a number of ways, as described above. More fundamentally, however, it provides pupils with the starting point for the development of their own criteria to judge themselves, as pupils and as persons. We use mirrors to see our reflections; that is to say, we use something outside ourselves that shows us what we are like. Reflexive self-assessment is similar. Pupils need to learn how to use their own 'mirrors' in religious education so that they can identify their own potential in their intellectual abilities, attitudes and values, and moreover to be proud of this potential. To achieve this poses some exciting challenges for education.

Appendix 1

Module traffic-light sheet

Shoah

Aim: To identify the key themes for the module.

SHOAH

Learning *about* Religion			
I understand religious and other explanations of evil	✓	✓	✓
I can describe how the holocaust happened	✗	✓	✓
I can describe *Yad Vashem*	✗	✗	✓
I know what the *Kindertransport* was.	✗	✗	✓
I know why the holocaust is a big thing for Jews.	–	✓	✓
Learning *from* Religion			
I can evaluate religious and other explanations of evil	–	✓	✓
I can empathise with Jewish experiences of persecution and escape	✗	✓	✓
I can relate the Jewish need to remember with my own experiences	✗	✓	✓
I can explain ~~why~~ that my views of good & evil are different to those of Jews.	✗	✓	✓

Appendix 2

Example of written self-assessment comments

Aim: To reflect on studying the holocaust.

Studying the holocaust does make me more tolerant because I understand how the jews feel about it

Module Review 31-3

Aim: To identify progress this term.

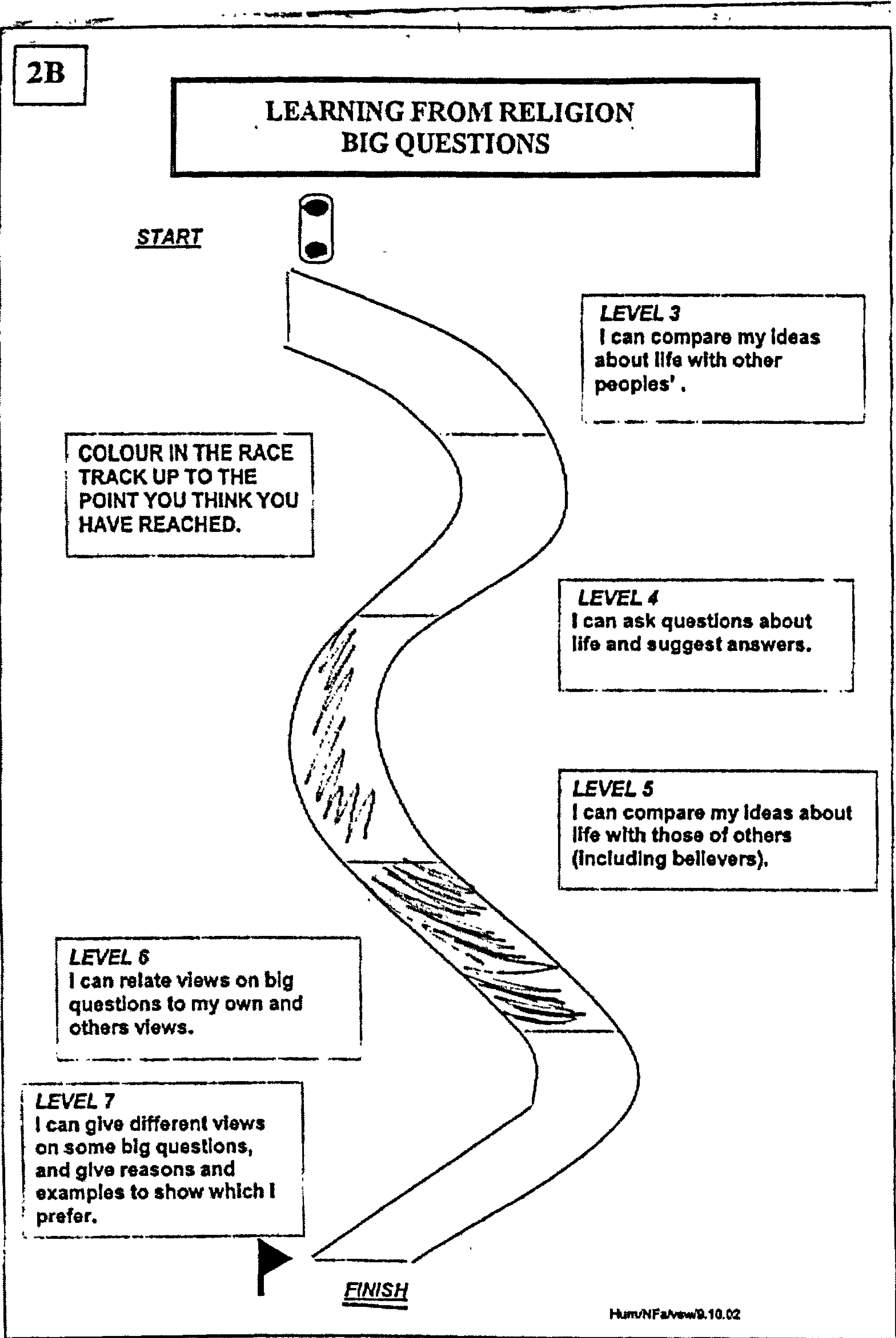
My knowledge has improved in all areas, mostly in Kindertransport as I didn't understand what that was at all. Know though I understand completely.

Studying the holocaust does make me more tolerant because I understand how the Jews feel about it...

My knowledge has improved in all areas, mostly in the Kindertransport as I didn't understand what that was at all. Know [read Now] though I understand completely.

Appendix 3

Level-path sheet



Appendix 4

Timetable for teaching and data collection in the main research

TERM 1: 'DOES GOD EXIST?' MODULE

- Start module with traffic lights sheet. Monitor at 4-week gaps, i.e. 4 dates in total. I choose most of the themes for this; pupils can identify two targets.
- Pupils to carry out reviews at end of reach lesson using a '*rapporteur*' – a student reporting back, focusing on 'learning from religion'.
- For imaginative work, e.g. on Paul's conversion, establish class criteria for assessment objectives. Then a peer-assessment exercise before comments only self-assessment.
- For second piece, e.g. 'does God exist?', use of levels in self-assessment. Do as peer-assessment exercise, then self-assessment, noting improvements.
- For review of all learning in the term, including the development of attitudes and values, conduct whole-class discussion.
- Conduct the first four interviews; start transcribing

TERM 2: 'EVIL AND SUFFERING' MODULE

- Start module with traffic lights sheet. Monitor at 4-week gaps, i.e. 4 dates in total. I choose most of the themes for this; they can identify two targets.

- Pupils to carry out reviews at end of reach lesson using a *'rapporteur'* – a student reporting back, focusing on 'learning from religion'.
- For imaginative piece, e.g. on escape on the *kindertransport*, class criteria for a peer-marking exercise.
- For evaluation exercise on explanations of evil, peer-assessment then self-assessment with comments only.
- For review of all learning in the term, including the development of attitudes and values, hold whole-class discussion
- Conduct the second four interviews; start transcribing

TERM 3: 'WEARING HIJAB' MODULE AND OPEN PROJECT

- Start module with traffic lights sheet. Monitor at 4-week gaps, i.e. 4 dates in total. I choose most of the themes for this; they can identify two targets.
- Pupils to carry out reviews at end of reach lesson using a *'rapporteur'* – a student reporting back, focusing on 'learning from religion'.
- For empathy exercise, e.g. hijab, discuss assessment objectives and conduct a peer-marking exercise.
- For evaluation exercise on whether schools should allow hijab, use levels; peer marking then self-marking.
- For review of all learning in the term, including the development of attitudes and values, whole class discussion
- Conduct the last four interviews; start transcribing

Appendix 5

Draft pilot questionnaire

Please tick the box that applies the most for each statement.

THIS TERM...	Agree a lot	Agree	Dis-agree	Dis-agree a lot
1. I understood the three different explanations of evil.				
2. I evaluated three different explanations of evil.				
3. I gave my opinions on evil.				
4. Self-assessment helps me learn.				
5. Doing the lesson review (ticking/crossing the Aim) helps me learn				
6. Using the Levelpath sheet helped me identify what I needed to do.				
7. Assessing the samples helped me to understand what a good piece of work was like.				
8. Assessing my own work helped me to understand what level I got.				
9. Assessing my own work helped me to understand what I need to do to get a higher level.				
10. I like knowing what level I am at.				
11. Using levels motivates me to do better.				

Appendix 6

Transcription conventions

Conventions from Torrance and Pryor (1998)

(*)	Inaudible (probably one word)
(**)	Inaudible phrase
(***)	Longer inaudible passage (e.g. sentence)
(*Tuesday)	Inaudible word, 'Tuesday' suggested by transcriber
-	short pause
disapp\	Incomplete word
these	Word emphasised
COME HERE	Words said very loudly in comparison to other utterances of this speaker
=	Rapid change of turn of speakers (used at the end of utterance and beginning of next speaker's utterance)
>It's mine<	Simultaneous speech
Italics	Non-textual material (stage directions)

Appendix 7

Pilot study codes

1. Traffic light process (lesson)
2. Traffic light process (module)
3. Levelpath sheet
4. Peer assessment using levels
5. Self-assessment using levels
6. Devising own criteria
7. Peer assessment to devise own criteria
8. Tolerance discussion
9. Social activity
10. Teaching implications
11. Self assessment styles/preferences
12. Feedback/feedforward
13. Motivation
14. Uncertainty of definitions

Appendix 8

Main research codes

1. SELF-ASSESSMENT
   a. FEEDBACK/FEEDFORWARD
   b. UNCERTAINTY OF DEFINITIONS
   c. TECHNIQUES
      i. feedback/feedforward
      ii. traffic light process (lesson)
      iii. traffic light process (module)
      iv. levelpath sheet
      v. peer assessment
      vi. self-assessment (levels)
      vii. devising own criteria
      viii. peer assessment to devise own criteria
      ix. tolerance discussion
      x. comparison across subjects
2. THE USE OF LEVELS
   a. Recognising 'place'
   b. Levels as 'target'
   c. Levels for 'marking'
   d. Timing of peer assessment
   e. Originality
   f. Others knowing your level
   g. Getting your own level wrong

3. ASSESSMENT AS SOCIAL ACTIVITY

a. Others knowing your level

b. Getting your own level wrong

c. Judging others wrongly – fear of condemnation

d. Judging others wrongly – deliberate spite

e. Shared construction

f. Embarrassment

g. Views of ‘the class’

h. Competition

4. UNDERSTANDING THE CRITERIA

a. Difficulties in knowing the criteria for self-assessment

b. Learning outcomes

c. Discussion

5. MOTIVATION

a. Motivated

b. Demotivated

6. LEARNING

a. Pupil’s image of self as a learner – general

i. Performance – successful

ii. Performance – unsuccessful

iii. Learning in general

b. Views of others

i. Brainy

ii. Stupid

c. Tests

7. LEARNING IN RE

a. Pupil's image of self as a learner – in RE

i. 'Learning about'

ii. Learning from'

iii. Wider aims of RE

b. Classroom activities

i. Discussion

ii. Empathy

iii. Writing

iv. Using whiteboard

v. Listening to others

vi. Respect

c. Wide views of the value of the subject

d. Education as an experiment

e. Behaviour

8. TEACHING IMPLICATIONS

a. 'Neutrality'

b. facilitator of discussion

c. teacher's character

d. lesson numbers

e. variety

f. as assessor

g. as facilitator of self-assessment

9. RESEARCH PROCESS

a. Reactivity

b. Exasperation

c. Social activity

d. Research questioning

e. Teacher researcher

    i. Over-Defensive

    ii. Behaviour management of the interview

    iii. Not pursuing certain questions

    iv. jokes

    v. teacher explanation

f. Group interview issues

    i. being the interviewer

    ii. bullying in the interview

g. Pupils' return to theme

Appendix 9

Main research interview timetable

GROUP	Interviewees	Interview 1	Interview 2	Interview 3
A	Helen Dave Chas David	15 November	24 February	20 July Chas absent; replaced by Penny
B	Ellie Jake Lucy Mike	15 November	1 April	7 July
C	Alan Aaron Lucy Sara	8 December	31 March	7 July
D	Billy Ahmed Kevin Annie	16 December	11 May Billy absent	4 July

Appendix 10

First pilot questionnaire

(n = 21)	Agree a lot	Agree	Dis-agree	Dis-agree a lot
1. I understood the three different explanations of evil.	10	11	0	0
2. I evaluated three different explanations of evil. (3 SPOILS)	8	8	2	0
3. I gave my opinions on evil.	9	10	2	0
4. Self-assessment helps me learn.	1	13	7	0
5. Doing the lesson review (ticking/crossing the Aim) helps me learn	3	9	9	0
6. Using the Levelpath sheet helped me identify what I needed to do. (4 SPOILS)	3	11	3	0
7. Assessing the samples helped me to understand what a good piece of work was like.	6	14	1	0
8. Assessing my own work helped me to understand what level I got. (1 SPOIL)	6	10	4	0
9. Assessing my own work helped me to understand what I need to do to get a higher level. (2 SPOIL)	9	9	1	0
10. I like knowing what level I am at.	12	9	0	0
11. Using levels motivates me to do better.	10	8	3	0

Appendix 11: Second pilot questionnaire

(n = 26)		Agree a lot	Agree	Dis-agree	Dis-agree a lot
I understood and evaluated the explanations of evil.	Boys	6	7		
	Girls	3	10		
	All	**9**	**17**		
I empathized with Jewish people involved in the *Kindertransport* and *Shoah*.	Boys	5	7	1	
	Girls	4	9		
	All	**9**	**16**	**1**	
I became more tolerant through studying the *Shoah*	Boys	3	7	3	
	Girls		7	5	1
	All	**3**	**14**	**8**	**1**
Generally, self-assessment helps me learn. 1 SPOILED	Boys	3	5	5	
	Girls	1	3	7	1
	All	**4**	**8**	**12**	**1**
Doing the lesson and module review (ticking/crossing the Aim and sheet) helps me learn.	Boys	4	6	3	
	Girls	2	7	3	1
	All	**6**	**13**	**6**	**1**
Using the levels on the essay helped me assess my work.	Boys	7	6		
	Girls	2	7	3	
	All	**9**	**13**	**3**	
Making our own standards from examples helped me assess my work.	Boys	3	6	4	
	Girls	3	5	5	
	All	**6**	**11**	**9**	
Discussing whether I have become more tolerant helped me assess my tolerance.	Boys	4	8	1	
	Girls	1	3	9	
	All	**5**	**11**	**10**	
Using different self-assessment activities is helpful. 2 SPOILED	Boys	1	8	4	
	Girls	2	6	2	1
	All	**3**	**14**	**6**	**1**
I have got better at assessing my own work in RE this term.	Boys	2	10	1	
	Girls	1	8	4	
	All	**3**	**18**	**5**	
I am better at identifying what I need to do to improve in RE. 1 SPOILED	Boys	1	9	3	
	Girls	1	6	5	
	All	**2**	**15**	**8**	
Pupils choosing the most helpful activity for self-assessment		**Tra-ffic light**	**Use Lev-els**	**Own crit-eria**	**Dis-cus-sion**
FIRST CHOICE	Boys		8	1	4
	Girls	4	7	1	1
	All	**4**	**15**	**2**	**5**
OVERALL CHOICE (Total weighted ranking: 1st = 4, 2nd = 3, etc.)	Boys	27	42	27	36
	Girls	38	42	23	27
	All	**65**	**84**	**50**	**63**

Appendix 12

Ten principles for assessment in religious education

1. **An assessment is an educational judgement, not necessarily a measurement.** The word 'assess' is derived from the Latin '*assidere*' which means 'to sit down beside' but can mean 'to besiege'. This policy should be interpreted in the light of the former, not the latter.

2. **Classroom assessment is essentially about working with pupils to help them make progress.** This is '*for* learning'. The issue is not where pupils have got to, but there they can get to, with assistance (see Assessment Reform Group (2002) *Assessment for learning: 10 Principles*). It is not primarily about school evaluation.

3. **The criteria should always be shared with the pupils.** This includes: module learning outcomes, lesson outcomes, level descriptors, group criteria (e.g. role plays), GCSE criteria, A level criteria, values and attitudes.

4. **The development of attitudes and values should also be assessed.** RE is partly on the curriculum because it nurtures tolerance and respect; these therefore need to be assessed too. This is only genuinely possible through self-assessment, but the teacher's role is to help adjust this reflexive capacity in pupils, while developing other forms of learning. Be aware of tensions between them.

5. **Everything that pupils do should be assessed thrice.** To achieve this, a range of strategies should be used, but especially self-assessment and peer assessment. For instance, if each lesson includes a plenary review of the lesson, and each module includes a plenary review of the module, then everything has been assessed twice. Peer assessment of activities such as role plays is also appropriate. This can also include the wider goal of learning in RE, such as attitudes and values. Teacher comments 'on the hoof' during learning are also part of this process.

6. **Formal teacher assessment is an essential part of this.** For some kinds of knowledge and certain skills, pupils will need teacher guidance on their work, through feedback. This should be *timely*, should identify strengths and make suggestions on how to improve, so as to 'close the gap'. Pupils should be encouraged to amend their work in the light of these suggestions; this should generally be at least twice a term.

7. **Occasionally, it will be appropriate to use levels or grades.** This should be no more than once a module at key stage 3 or once a term at KS 4 because pupils may become too ego-focused. These should be used *with* comments, and pupils must be given the opportunity to improve the particular piece of work, and therefore the level or grade. Avoid rewards for quality alone. These levels or grades will inform our monitoring and reporting. At A level, grades should be more frequent.

8. **Teachers should monitor their pupils' ability to self-assess.** Some pupils will be better at making sense of their progress than others: more task-

focused, less ego-focused. This will often overlap with pupils' ability to 'learn from religion' or evaluation for GCSE. Assessment therefore needs to be differentiated in order to share control of learning.

9. **Pupils' attitudes to assessment in RE will be coloured by their experiences elsewhere in school.** This will affect their attitudes to formative and summative techniques. They may feel judged by the use of levels; they may not know what to do with comments. Just because you intend levels or grades to be formative does not mean the pupils treat them in this way. You know what they are like near exams.

10. **Pupils should learn to value their values, not just to value what exam boards can measure.** If they do this they will value RE, and value appropriately what exam boards *can* measure.

References

Abu-Lughod, L. (1988) *Veiled Sentiments: Honor and Poetry in a Bedouin Society* (Berkley: University of California)

Afdal, G. (2006) *Tolerance and Curriculum: conceptions of tolerance in the multi-cultural unitary school Norwegian compulsory school* (Münster: Waxmann)

Afdal, G. (2007a) 'Globalization and research language in religious education' paper presented at European Network for RE through Contextual Approaches conference, Utrecht, March 2007

Afdal, G. (2007b) 'Participative research in religious education: an argument for a dialectical understanding of theory and practice' in C. Bakker and H-G. Heimbrock *Researching RE teachers, RE teachers as researchers* (Münster: Waxmann)

Afdal, G. (forthcoming) 'Religious Education as a Research Discipline: An Activity Theoretical Perspective' in *British Journal of Religious Education*

Aikenhead, G. (1997) 'A framework for reflecting on assessment and evaluation', in *Globalization of science education – papers from the Seoul international conference,* 195-199 (Seoul: Korean Educational Development Institute). Also at: http://www.usask.ca [downloaded 30/07/07, page references to the downloaded text]

Alexander, R. (2005) 'Culture, dialogue and learning: notes on an emerging pedagogy' keynote paper at International Association for Cognitive Education

and Psychology (IACEP) 10th International Conference, University of Durham, UK, 10-14 July 2005

Allal, L. (2007) 'Régulations des apprentissages : orientations conceptuelles pour la recherche et la pratique en éducation' in L. Allal and L. Mottier Lopez (eds.) *Régulation des apprentissages en situation scolaire et en formation* (Brussels: de Boeck)

Allal, L. and Mottier Lopez, L. (2005) 'Formative Assessment of Learning: a review of publications in French' in OECD (ed.) *Formative Assessment: Improving Learning in Secondary Classrooms* (Paris: OECD)

Allal, L. and Mottier Lopez, L. (eds.) (2007) *Régulation des apprentissages en situation scolaire et en formation* (Brussels: de Boeck)

Altricher, H., Posch, P. and Somekh, B. (1993) *Teachers Investigate their Work: An Introduction to the Methods of Action Research* (London: Routledge)

Anscombe, E. (1958) 'Modern Moral Philosophy' in *Philosophy* 33, 1-19

Aristotle (1953) *The Ethics of Aristotle* (London: Penguin)

Arweck, E. and Stringer, M. (eds.) (2002) *Theorising Faith: The Insider/Outsider Problem in the Study of Ritual* (Birmingham: University of Birmingham Press)

Asher, H. (1984) 'The research process' in H. Asher, H. Weisberg, J. Kessel, and W. Shively (eds.) *Theory-Building and data Analysis in the Social Sciences* (Knoxville: University of Tennessee Press)

Assessment Reform Group (2002) *Assessment for Learning: 10 Principles* (London: Assessment Reform Group)

Association for Achievement and Improvement through Assessment (2002) *Secondary Assessment Practice: Self Evaluation and Development Materials* (London: AAIA)

Association for Achievement and Improvement through Assessment, North East Region (2003) *Self-Assessment* (London: AAIA)

Association of Religious Education Inspectors, Advisors and Consultants (1998) *Towards National Standards in Religious Education* (self-published)

Astley, J., Francis, L., Burton, L. and Wilcox, C. (1997) 'Distinguishing between Aims and Methods in RE: a Study among Secondary School Teachers' in *British Journal of Religious Education* 19:3, 171-184

Attfield, D. (1996) 'Learning from Religion' in *British Journal of Religious Education* 18:2 78-84

Attfield, D. (1997) 'A response to Robert Ellis' in *British Journal of Religious Education* 19:3, 133

Bakker, C. and Heimbrock, H-G. (2007) 'Religious Education teachers: Reflective Practitioners and Researchers' in C. Bakker and H-G. Heimbrock (eds.) *Researching RE teachers, RE teachers as researchers* (Münster: Waxmann), 7-14

Barry, C. (1998) 'Choosing Qualitative Data Analysis Software: Atlas/ti and Nudist Compared' in *Sociological Research Online,* 3:3 http://www.socresonline.org.uk/socresonline/3/3/4.html [downloaded 10/04/06]

Barthes, R. (1970) *S/Z* (Paris: Seuil)

Barthes, R. (1977) 'The Death of the Author' in *Image Music Text* (ed. S. Heath) (London: Fontana)

Baumfield, V., Bowness, C., Cush, D. and Miller, J. (1993) *A Third Perspective* (Exeter: University of Exeter)

Baumfield, V., Bowness, C., Cush, D. and Miller, J. (1994) 'Model Syllabuses: the debate continues' in *REsource* 18:2, 3-6

Beck, J. (1979) *Handbook in Social Philosophy* (New York: Macmillan)

Bicknell, E. (1919) *A Theological Introduction to the Thirty-Nine Articles* (London: Longmans)

Biggs, J. (1998) 'Assessment and classroom learning: a role for summative assessment' in *Assessment in Education* 5:1, 103-110

Black, P. (1998) *Testing: Friend or Foe?* (London: Falmer)

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2002) *Working inside the black box: Assessment for learning in the classroom* (London: School of Education, King's College)

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2003) *Assessment for Learning: Putting it into Practice* (Maidenhead: Open University Press)

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2005) 'Dissemination and evaluation: a response to Smith and Gorard' in *Research intelligence,* November 2005, issue 93, 3-4

Black, P., McCormick, R., James, M. and Pedder, D. (2006) 'Learning How to Learn and Assessment for Learning: a theoretical inquiry' in *Research Papers in Education* 21:2, 119 - 132

Black, P. and Wiliam, D. (1998a) 'Assessment and classroom learning' in *Assessment in Education* 5:1, 7-74

Black, P. and Wiliam, D. (1998b) *Inside the black box: raising standards through classroom assessment* (London: School of Education, King's College)

Black, P. and Wiliam, D. (2006) 'Assessment for learning in the Classroom' in Gardner, J. (ed.) *Assessment and Learning* (London: Sage), 9-26

Blandin, A., Borne, D., Legrand, A., Le Guillou, P., Sherringham, M. and Willaime, J-P. (2004) 'La religion à l'école en France' in *Revue Internationale d'éducation: dossier – école et religion,* No. 36, July 2004

Blaylock, L. (2000) 'Issues in achievement and assessment in religious education in England: Which way should we turn?' in *British Journal of Religious Education* 23:1, 45-58

Blaylock, L. (2003) 'Religious education and citizenship: some reflections' in R. Jackson, *International perspectives on Citizenship, Education and Religious Diversity* (London: Routledgefalmer)

Blaylock, L. (2007) personal communication

Bloom, B. (1956) *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain* (New York: David McKay Co Inc.)

Boud, D. (1986) *Implementing Student Self Assessment* (Sydney: HERDSA)

Bourdieu, P. (1977) *Outline of a Theory of Practice* (Cambridge, Cambridge University Press).

Bourdieu, P. and Wacquant, L. (1992) *An Invitation to Reflexive Sociology* (Cambridge: Polity)

Bowe, R. and Ball, S. with Gold, A. (1992) *Reforming Education and Changing Schools* (London: Routledge)

British Broadcasting Corporation (2002) 'Xenophobia slur on history lessons' at news.bbc.co.uk/1/hi/education/2559011.stm [downloaded 24/03/2008]

British Broadcasting Corporation (2004) 'Blair praises exams "successes"' at news.bbc.co.uk/1/hi/education/ 3634768.stm [downloaded 29/11/2006]

British Educational Research Council (2003) 'Ethical Guidelines for educational

research: consultation of members' in *Research Intelligence* 82 (Southwell: BERA)

Bronfenbrenner, U. (1977) 'Towards an experimental ecology of human development' in *American Psychologist* 32, 513-531

Brooks, V. (2002) *Assessment in Secondary Schools* (Buckingham: Open University Press)

Brooks, V. (2006) 'The Role of External Examinations in the Making of Secondary Modern Schools in England 1945–65' in *History of Education*, 2006, 1–21

Buckle, C. (1990) 'The Types, Purposes and Effects of Assessments' in R. Riding and S. Butterfield (eds.) *Assessment and Examination in the Secondary School* (London: Routledge), 60-79

Burn, J. and Hart, C. (1988) *The Crisis in Religious Education* (London: Educational Research Trust)

Burrell, G. and Morgan, G. (1979) *Sociological Paradigms and Organization Analysis* (London: Heinemann)

Burton, D. and Bartlett, S. (2005) *Practitioner Research for Teachers* (London: Paul Chapman Publishing)

Butler, R. (1987) 'Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest and

performance' in *Journal of Educational Psychology* 79, 474-482

Butler, R. (1988) 'Enhancing and Undermining Intrinsic Motivation: the Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance' in *British Journal of Educational Psychology* 58, 1-14

Campanale, F. (2007) 'Evaluation réflexive en formation professionnel et évaluation interactive dans les classes' in L. Allal and L. Mottier Lopez (eds.) *Régulation des apprentissages en situation scolaire et en formation* (Brussels : de Boeck), 191-206

Campbell, A., McNamara, O., and Gilroy, P. (2004) *Practitioner Research and Professional Development in Education* (London: Paul Chapman Publishing)

Campbell, E. (2003) *The Ethical Teacher* (Maidenhead: Open University Press)

Carr, D. (2006) 'Professional and personal values and virtues in education and teaching' in *Oxford Review of Education* 32:2, 171-183

Carr, W. and Kemmis, S. (1983) *Becoming Critical: Knowing through Action Research* (Geelong: Deakin University Press)

Carr, W. and Kemmis, S. (2005) 'Staying critical' in *Educational Action Research* 13:3, 347-358

Cohen, L., Manion, L. and Morrison, K. (2000) *Research Methods in Education* (5th edition) (London and New York: RoutledgeFalmer)

Comte-Sponville, A. (2003) [trans. C. Temerson] *A Short Treatise on the Great*

Virtues: The Uses of Philosophy in Everyday Life (London: Vintage) [French Edition, *Petit Traité des Grandes Vertus* (1996)]

Copley, T. (1997) *Teaching Religion: Fifty years of religious education in England and Wales* (Exeter: University of Exeter Press)

Copley, T. (2000) *Spiritual Development in the State School* (Exeter: University of Exeter Press)

Copley, T. (2005) *Indoctrination, Education and God: the Struggle for the Mind* (London: SPCK)

Copley, T., Priestley, J., Wadman, D. and Coddington, V. (1991) *Forms of Assessment in Religious Education: The main report of the FARE Project* (Exeter: FARE Project)

Council of Europe (2007) 'Draft recommendation on the religious dimension of intercultural education: principles, objectives and teaching approaches' unpublished document (Strasbourg: Council of Europe)

Cowie, B. (2002) *Formative assessment in science classrooms* Unpublished PhD thesis, University of Waikato, New Zealand

Cowie, B. (2005) 'Pupil commentary on assessment for learning' in *The Curriculum Journal* 16:2, 137-151

Cox, C. (1988) 'Foreword', in J. Burn and C. Hart (eds.) *The Crisis in Religious Education* (London: The Educational Research Trust)

Crooks, T.J. (1988) 'The impact of classroom evaluation practices on students' in *Review of Educational Research* 58, 438-481

Culler, J. (1983) *Barthes* (London: Fontana)

Daugherty, R. (2007) 'Mediating academic research: the Assessment Reform Group experience' in *Research Papers in Education* 22:2, 139-153

Daugherty, R. and Ecclestone, K. (2006) 'Constructing Assessment for learning in the UK Policy Environment' in J. Gardner (ed.) *Assessment and Learning* (London: Sage), 149-168

Davis, A. (1999) *Educational assessment: a critique of current policy* (London: Philosophy of Education Society of Great Britain)

Davis, B. (1997) 'Listening for differences: an evolving conception of mathematics teaching' in *Journal for Research in Mathematics Education* 28, 355-376

Day, J. and Cordon, L. (1993) 'Static and dynamic measures of ability; an experimental comparison' in *Journal of Educational Psychology* 85, 76-82

Deakin Crick, R., McCombs, B., Haddon, A., Broadfoot, P., and Tew, M. (2007) 'The ecology of learning: factors contributing to leaner-centred classroom cultures' in *Research Papers in Education* 22:3, 267-307

Dearing, R. (1994) *The National Curriculum and its Assessment* (London: BERA

and Paul Chapman)

Debray, J. (1992) *L'enseignement du fait religieux dans l'école publique : rapport au ministre de l'Education Nationale* (Paris: Odile Jacob)

Deng, F. M. (1980) *Dinka Cosmology.* (London: Ithaca Press)

Department for Education and Skills (2003) *Key Messages from the Key Stage 3 National Strategy* (London: DfES) [Reference: DfES 0208/2003]

DePaul, M. and Zagzebski, L. (eds.) (2003) *Intellectual Virtue: Perspectives from Ethics and Epistemology* (Oxford: Oxford University Press)

Devon Local Education Authority (undated) *Exemplification of RE Work* (www.devon.gov.uk) [downloaded 30/01/2004]

Draycott, P. (2006) *Assessed RE* (Derby: CEM)

Duman, D. (2004) 'Laïcité et islamisme en Turquie aujourd'hui' in *Revue Internationale d'éducation: dossier – école et religion*, No. 36, July 2004

Durkheim, E. (1937) *Les Règles de la Méthode Sociologique* (Paris: Presses Universitaires de France) [The Rules of Sociological Method]

Dweck, C. (1999) *Self-Theories: their Role in Motivation, Personality and Development* (London: Psychology Press)

Ebel, R. and Frisbie, D. (1991) *Essentials of Educational Measurement*

(Eaglewood Cliffs: Prentice Hall)

Ecclestone, K. (2007) 'Commitment, compliance and comfort zones: the effects of formative assessment on vocational education students' learning careers' in *Assessment in Education: Principles, Policy & Practice* 14:3, 315 – 333

Ecclestone, K. and Pryor, J. (2003) '"Learning careers" or "Assessment careers"? The impact of assessment systems on learning' in *British Educational Research Journal* 29:4, 471-488

Economides, K. (ed.) (1998) *Ethical Challenges to Legal Education and Conduct* (Oxford: Hart)

Elliott, J. (1991) *Action Research for Educational Change* (Milton Keynes: Open University Press)

Elliott, J. (2005) 'Becoming Critical: the failure to connect' in *Educational Action Research* 13:3, 359-373

Elliott, J. (2007) 'Assessing the quality of action research' in *Research Papers in Education* 22:2, 229-246

Ellis, R. (1997) 'Revelation, wisdom, and learning from religion: A response to D.G. Attfield' in *British Journal of Religious Education* 19:2, 95-103

Ellwood, J. (2006) 'Formative assessment: possibilities, boundaries and limitations' in *Assessment in Education* 13:2, 215-232

Ely, M. with Anzul, M., Freidman, T., Garner, D. and McCormack Steinmetz, A. (1991) *Doing Qualitative Research: Circles within Circles* (London: Falmer)

Endicott, L. (2001) *Activity Booklet 1: ethical sensitivity* (Notre Dame: University of Notre Dame)

Engeström, Y. (2005). *Developmental work research. Expanding activity theory in practice* (Berlin: Lehmanns Media)

Erricker, C. and Erricker, J. (2000) *Reconstructing Religious, Spiritual and Moral Education* (London: RoutledgeFalmer)

Eraut, M. (1994) *Developing Professional Knowledge and Competence* (London: Falmer)

Estivalèzes, M. (2005) *Les religions dans l'enseignement laïque* [Religions in *laïque* teaching] (Paris: Presses Universitaires de France)

Estivalèzes, M. (2007) 'Teaching about religions in a *laïc* educational system' in C. Bakker, and H-G. Heimbrock (eds.) *Researching RE teachers, RE teachers researching* (Münster: Waxmann)

Evans-Pritchard, E. (1940) *The Nuer* (Oxford: Oxford University Press)

Everington, J. (2000) 'Mission impossible? Religious education in the 1990's', in M. Leicester, C, Mogdil & S. Mogdil (eds.) *Education, Culture and Values, Volume 5: Spiritual and religious education* (London, Falmer Press), 183–196

Everington, J. (2007) 'Freedom and direction in religious education: the case of English trainee teachers and "learning from religion"' in C. Bakker and H-G Heimbrock (eds.) *Researching RE teachers, RE teachers as researchers* (Münster: Waxmann), 111-124

Fancourt, N. (2003a) *Assessing Empathy* unpublished Farmington paper [www.farmington.ac.uk]

Fancourt, N. (2003b) 'Whole class strategies in religious education' in *Support for Learning* 18:3, 112-116

Fancourt, N. (2005a) 'Challenges to Self-Assessment in Religious Education' in *British Journal of Religious Education* 27:2, 115-126

Fancourt, N. (2005b) 'Review of Campbell, Anne, McNamara, Olwen, and Gilroy, Peter (2004) *Practitioner Research and Professional Development in Education* (London: Paul Chapman Publishing)', in *History of Education and Administration* 37:1, 113-115

Fancourt, N. (2006) 'Self-assessment for beginners' in *REToday*, Spring 2006, Vol. 23:2, 28-29

Fancourt, N. (2007) 'The "dialogical" teacher: should teachers express their commitments in the classroom?' in C. Bakker and H-G. Heimbrock *Researching RE teachers, RE teachers as researchers* (Münster: Waxmann), 53-70

Fancourt, N. (2008) 'Review of Teaching Religious Education by Julian Stern' in *British Journal of Religious Education*, 30:1, 83-85

Felderhof, M. (2004) 'The New National Framework for RE in England and Wales: a critique' in *Journal of Beliefs and Values* 25:2, 241-148

Fielding, M. (2007) 'Jean Rudduck (1937-2007) "Carving a new order of experience": a preliminary appreciation of the work of Jean Rudduck in the field of student voice' in *Educational Action Research* 15:3, 323-336

Filer, A. and Pollard, A. (2000) *The Social World of Pupil Assessment: the Processes and Contexts of Primary Schooling* (London and New York: Continuum)

Flutter, J. and Rudduck, J. (2004) *Consulting pupils: what's in it for schools?* (London: RoutledgeFalmer)

Fontana, D. and Fernandez, M. (1994) 'Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils' in *British Journal of Educational Psychology* 64, 407-417

Forde, D. (1954) *African Worlds: Studies in the Cosmological Ideas and Social Values of African Peoples* (Oxford: International African Institute/Oxford University Press)

Francis, L. and Fearn, M. (2001) 'Assessing and Interpreting Learning preferences among A-level Religious Studies Students: the Revised Index of Learning Styles' in *British Journal of Religious Education* 24:1 41-51

Freathy, R. (2008) 'Three Perspectives on Religious Education and Education for

Citizenship in English Schools, 1934-1944: Cyril Norwood, Ernest Simon and William Temple' in *British Journal of Religious Education* 30:2, 103-113

Fredriksen, J. and White, B. (1997) 'Reflective assessment of students' research within an inquiry-based middle school science curriculum', in *Proceedings of the Annual Meeting of the AERA, Chicago, 1997*

Fuller, R. and Petch, A. (1995) *Practitioner Research: the reflexive social worker* (Buckingham: Open University Press)

Furlong, J., Barton, L., Miles, S., Whiting, C. and Whitty, G. (2000) *Teacher education in transition: re-forming professionalism?* (Buckingham: Open University Press)

Gardner, J. (2006) 'Assessment and Learning: an Introduction' in J. Gardner (ed.) *Assessment and Learning* (London: Sage), 1-5

Gardner, P. (1992) 'Propositional attitudes and multicultural education or believing that others are mistaken' in J. Horton and P. Nicholson (eds.) *Toleration: philosophy and practice* (Aldershot: Avebury), 67-90

Gennerich, C. (2007) Interreligious Styles of Communication in German Religious Education Teachers' Perspectives – a critical Counterpart to Theories in religious Education? in C. Bakker and H-G. Heimbrock (eds.) *Researching RE teachers, RE teachers researching* (Münster: Waxmann), 235-250

Giddens, A. (1993) *New Rules of Sociological Method* (2nd Edition) (Cambridge: Polity Press)

Gipps, C. (1994) *Beyond Testing: Towards a Theory of Educational Assessment* (London: Hodder and Stoughton)

Gipps, C. and Murphy, P. (1994) *A Fair Test? Assessment, Achievement and Equity* (Buckingham: Open University Press)

Glaser, B. and Strauss, A. (1967) *The Discovery of Grounded Theory* (Chicago: Aldine)

Goodson, I. (1990) 'Nations at Risk and National Curriculum: Ideology and Identity', in *Politics of Education Association Yearbook* (London: Taylor and Francis), 219-252

Gorard, S. and Taylor, C. (2004) 'What is 'triangulation'?' in *Building Research Capacity*, February 2004, 7-9

Gough, A. and Sharpley, B. (2005) 'Toward effective teaching and learning: stories of primary schools' environmental science interest and action' in *Educational Action Research* 13:2, 191-212

Greco, J. (2004) 'Virtue Epistemology' in *Stanford Encyclopedia of Philosophy* on line at: plato.stanford.edu/entries/epistemology-virtue [downloaded 26/04/2005]

Griffiths, M. (1998) *Educational research for Social Justice: Getting off the Fence* (Buckingham: Open University Press)

Grimmitt, M. (1981) 'When is Commitment a Problem in Religious Education?'

British Journal of Educational Research, 29:1, 42-53

Grimmitt, M. (1982) 'World religions and personal development' in R. Jackson (ed.) *Approaching World Religions* (London: John Murray), 136-149

Grimmitt, M. (1987) *Religious Education and Human Development: the relationship between studying religions and personal, social and moral education* (Essex: McCrimmons)

Grimmitt, M. (1991) 'The use of religious phenomena in schools: some theoretical and practical considerations' in *British Journal of Religious Education*, 13:2, 77-88

Grimmitt, M. (2000) 'The Captivity and Liberation of Religious Education and the Meaning and Significance of Pedagogy' in M. Grimmitt (ed.), *Pedagogies of Religious Education* (Great Wakering: McCrimmonds), 7-23

Gronlund, N. (1985) *Measurement and Evaluation in Teaching* (New York: Macmillan)

Haakedal, E. (2007) 'Situated Practice among Religious Education Teachers. A Discussion of School Rituals, Cultural Contexts and Professional Ethics' in C. Bakker and H-G. Heimbrock *Researching RE teachers, RE teachers as researchers* (Münster: Waxmann), 37-52

Habermas, J. (1972) *Knowledge and Human interests* [tr. J. Shapiro] London: Heinemann)

Halai, A. (2005) 'Action research to study impact: is it possible?' in *Educational Action Research* 12:4, 515-534

Hammersley, M. (2004) 'Action research: a contradiction in terms?' in *Oxford Review of Education* 30:2, 165-181

Hargreaves, A. (1997) 'Restructuring Restructuring: Postmodernity and the Prospects for Educational Change' in A. Halsey, H. Lauder, P. Brown and A. Stuart Wells (eds.) *Education: Culture, Economy, Society* (Oxford: Oxford University Press), 338-353

Harlen, W. (2006) 'On the inter-relationship between Assessment for Formative and Summative Purposes' in J. Gardner (ed.) *Assessment and Learning* (London: Sage), 103-117

Harlen, W. and Deakin Crick, R. (2002) 'A systematic review of the impact of summative assessment and tests on students' motivation for learning' in *Research Evidence in Education Library.* (London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London) [Downloaded from eppi.ioe.ac.uk on 9/02/2006]

Harris, R. and Foreman-Peck, L. (2001) 'Learning to Teach History Writing: discovering what works' in *Educational Action Research* 9:1, 97-109

Hardy, A. (1979) *The Spiritual Nature of Man* (Clarendon Press: Oxford)

Harrap's Shorter Dictionary: French-English (1982) (London/Paris: Harrap Books)

Hattie, J. Biggs, J. and Purdie, N (1996) 'Effects of learning skills interventions on student learning: a meta-analysis' in *Review of Educational Research*, Vol 66, 963-136

Hay, D. (1982) *Exploring Inner Space* (Penguin: Harmonsworth)

Hay, D. (2000) 'The Religious Experience and Education Project: Experiential Learning in Religious Education' in M. Grimmitt (ed.) *Pedagogies of Religious Education* (Great Wakering: McCrimmons), 70-87

Hayward, Mary (2007) *Christianity in Religious Education at Key Stage 3*, Warwick Religions and Education Research Unit Occasional Paper 4, (Coventry: WRERU)

Herzog, T. (1996) *Research Methods in the Social Sciences* (New York: Harper Collins)

Hick, J. (1983) *The Second Christianity* (London: SCM Press)

Honderich, T. (ed.) (1995) *The Oxford Companion to Philosophy* (Oxford: Oxford University Press)

Hookway, C. (2003) 'How to be a Virtue Epistemologist' in M. DePaul and L. Zagzebski (eds.) *Intellectual Virtue: Perspectives from Ethics and Epistemology* (Oxford: Oxford University Press), 183-202

Hopwood, N. (2007) 'Values in geographic education: the challenge of attending to learners' perspectives' in *Oxford Journal of Education,* 1-20 (iFirst Article)

Hull, J. (1994) 'Editorial: Model Syllabuses' in *British Journal of Religious Education* Vol. 17:1, 2-4

Hull, J. (1995) 'Religion as a series of religions: a comment on the SCAA Model Syllabuses', in SHAP (ed.) *From syllabus to schemes: planning and teaching RE* (London, Shap Working Party).

Illich, I. (1973) *Deschooling Society* (London: Penguin Books)

Ions, E. (1977) *Against Behaviouralism: a Critique of Behavioural Science* (Oxford: Basil Blackwell)

Ipgrave, J. (1999) 'Issues in the Delivery of Religious Education to Muslim Pupils: Perspectives from the Classroom' in *British Journal of Religious Education* Vol. 21:3, 146-157

Ipgrave, J. (2001) *Pupil-to pupil dialogue in the Classroom as a Tool for Religious Education,* Warwick Religions and Education Research Unit Working Papers 2 (Coventry: University of Warwick)

Ipgrave, J. (2002) 'Inter faith Encounter and Religious Understanding in an Inner City Primary School' unpublished PhD thesis, University of Warwick

Ipgrave, J. (2003) 'Dialogue, citizenship and religious education' in R. Jackson (ed.) (2003) *International perspectives on Citizenship, Education and Religious Diversity* 147-168 (London: RoutledgeFalmer)

Ipgrave, J. and MacKenna, U. (2007) 'Values and Purposes: teacher perspectives on the "Building E-bridges" Project for Inter Faith dialogue between Children across the UK' in C. Bakker and H-G Heimbrock *Researching RE teachers, RE teachers researching* (Münster: Waxmann), 215-234

Jackson, R. (1993) 'Religious Education and the Arts of Interpretation' in D. Starkings (ed.) *Religion and the Arts in Education: Dimensions of Spirituality* (London: Hodder and Stoughton) 157-166

Jackson, R. (1997) *Religious Education: An Interpretive Approach* (London: Hodder and Stoughton)

Jackson, R. (2000a) 'Law, Politics and Religious Education in England and Wales: some History, some Stories and some Observations' in M. Leicester, C, Mogdil & S. Mogdil (eds.) *Education, Culture and values, Volume 5: Spiritual and religious education* (London, Falmer Press), 86–98

Jackson, R. (2000b) 'The Warwick Religious Education Project: The Interpretive Approach to Religious Education', in M. Grimmitt (ed.) *Pedagogies of Religious Education* (Great Wakering: McCrimmonds), 130-152

Jackson, R. (ed.) (2003) *International Perspectives on Citizenship, Education and Religious Diversity* (London: RoutledgeFalmer)

Jackson, R. (2004a) 'Editorial' in *British Journal of Religious Education* 26:3, 1

Jackson, R. (2004b) *Rethinking Religious Education and Plurality: Issues in Diversity and Pedagogy* (London: Routledge Falmer)

Jackson, R. (2005) 'Intercultural education, religious plurality and teaching for tolerance: interpretive and dialogical approaches' in R. Jackson and U. McKenna (ed.) *Intercultural Education and Religious Plurality* (Oslo: The Oslo Coalition on Freedom of Religion and Belief) 5-14

Jackson, R. (2008) 'Contextual religious education and the interpretive approach' in *British Journal of Religious Education* 30:1, 13-24

Jackson, R. and Fujiwara, S. (2007) 'Towards religious education for peace' in *British Journal of Religious Education* 29:1, 1-15

Jackson, R., Miedema, S., Weisse, W. and Willaime, J-P. (eds.) (2007) *Religion and Education in Europe: Developments, Contexts and Debates* (Münster: Waxmann)

Jackson, R. and Nesbitt, E. (1992) 'The diversity of experience in the religious upbringing of children from Christian families in Britain' in *British Journal of Religious Education* 15:1, 19-28

Jackson, R. and O'Grady, K. (2007) 'The Warwick REDCO Community of Practice' unpublished paper

James, M. (2006) 'Assessment, teaching and theories of learning' in J. Gardner (ed.) *Assessment and Learning* (London: Sage), 47-60

James, M., McCormick, R., Black, P., Carmichael, P., Drummond, M-J., Fox, A., MacBeath, J., Marshall, B., Pedder, D., Procter, R., Swaffield, S., Swann, J. and

Wiliam, D. (2007) *Improving Learning How to Learn: Classrooms, Schools and Networks* (London: Routledge)

James, M. and Pedder, D. (2006) 'Professional learning as a condition for assessment for learning' in J. Gardner (ed.) *Assessment and Learning* (London: Sage), 27-43

Kalve, P. (1996) 'Some Aspects of the work of Michael Grimmitt' in *British Journal of Religious Education* 18:3, 181-190

Kay, W. (2005) 'A Non-statutory framework for religious education: issues and opportunities' in *British Journal of Religious Education* 27:1, 41-52

Keast, J. (2003a) 'Issues in the Teaching of Religious Education: Assessing Achievement in RE from early years to A level', in L. Broadbent and A. Brown (eds) *Issues in Religious Education* (London: RoutledgeFalmer), 56-70

Keast, J. (2003b) Keynote speech at HMI RE Conference, Southwark Cathedral, March 2003

Kemmis, S. (1993) 'Action Research' in M. Hammersley (ed.) *Education Research: Current Issues* (London: Paul Chapman Publishing)

Kemmis, S. (2006) 'Participatory Action Research and the Public Sphere' in *Educational Action Research* 14:4, 459-476

Kincaid, M. (1991) *How to Improve Learning in RE* (London: Hodder and Stoughton)

Kluger, A. and DeNisi, A. (1996) 'The effects of feedback interventions on performance: a historical review, meta-analysis, and a preliminary feedback Intervention Theory' in *Psychological Bulletin*, 119(2), 254-284

Kumar, R. (2005) *Research methodology: a step-by-step guide for beginners* (London: Sage)

Laveault, D. (2007) 'De la «régulation» au «réglage» : élaboration d'un modèle d'autoévaluation des apprentissages' in L. Allal and L. Mottier Lopez (eds.) *Régulation des apprentissages en situation scolaire et en formation* (Brussels : de Boeck), 205-234

Learn how to Learn (2002) 'AfL in action: Peer and Self Assessment' published on www.learntolearn.ac.uk [downloaded 28/09/07]

Leitch, R., Gardner, J., Mitchell, S., Lundy, L., Odena, O., Galanouli, D. and Clough, P. (2007) 'Consulting Pupils in Assessment for Learning classrooms: the twists and turns of working with students as co-researchers' in *Educational Action Research* 15:3, 459-478

Lewin K. *Resolving Social Conflicts* (NewYork: Harper)

Lewins, A. and Silver, C. (2005) 'Choosing a CAQDAS Package: a working paper' on http://caqdas.soc.surrey.ac.uk/ChoosingLewins&SilverV3Nov05.pdf [downloaded 14/04/06]

Lewis, A. (1992) 'Group child interviews as a research tool' in *British Educational*

Research Journal, Vol. 18, No.4, 413-21

Lienhardt, G. (1961) *Divinity and Experience: the Religion of the Dinka* (Oxford: Oxford University Press)

Lienhardt, G. (1964) *Social Anthropology* (Oxford: Oxford University Press)

Looney, J. and Wiliam, D. (2005) 'Implementing Formative Assessment in a High stakes Environment' in Organisation for Economic Co-operation and Development (ed.) *Formative Assessment: Improving Learning in Secondary Classrooms* (Paris: OECD), 129-148

McCormick, R. and James, M. (1988) *Curriculum Evaluation in Schools* (London: Croom Helm)

McDonald, B. and Boud, D. (2003) 'The Impact of Self-Assessment on Achievement: the Effects of Self-Assessment Training on Performance in External Exams' in *Assessment in Education* 10:2, 209-220

McFadden, M. and Munns, G. (2002) 'Student Engagement and the Social Relations of Pedagogy' in *British Journal of Sociology of Education*, 23(3), 357-366

MacIntyre, A. (1981) *After Virtue: a study in moral theory* (London: Duckworth)

McKinnon, C. (2003) 'Knowing Cognitive Selves' in DePaul, M. and Zagzebski, L. (eds.) *Intellectual Virtue: Perspectives from Ethics and Epistemology* (Oxford: Oxford University Press), 227-254

Maitles, H. Cowan, P. and Butler, E. (2006) 'Never Again! Does Holocaust Education Have an Effect on Pupils' Citizenship Values and Attitudes?' at http://www.scotland.gov.uk/Publications/2006/09/06133626/0 [downloaded 24/03/2008]

Malinowski, B. (1982) 'The Diary of an Anthropologist' in R.G. Burgess (ed.) *Field research: a Sourcebook and Field Manual* (London: George Allen and Unwin), 200-205

Maqsud, M. and Pillai, C. M. (1991) 'Effects of self-scoring on subsequent performances in academic achievement tests' in *Educational Research* 33, 151-154

Marchall, B. and Drummond M-J. (2006) 'How teachers engage with assessment for learning: lessons from the classroom' in *Research Papers in Education* 21:2, 133-149

Marquez-Zenkov, K., Harmon, J., van Lier, P., and Marquez-Zenkov, M. (2007) 'If they listen to us about life, we'll listen to them about school: seeing city students' ideas about "quality" teachers' in *Educational Action Research* 15:3, 403-415

Marx, K. (1956) *Selected Writings in Sociology and Social Philosophy* (London: Penguin)

Maxwell, J. (1992) 'Understanding and validity in qualitative research' in *Harvard Educational Review* 62:3, 279-300

Maybury, J. and Teece, G. (2005) 'Learning from what? A question of subject focus in religious education in England and Wales' in *Journal of Beliefs and Values*, 27:2, 179-190

Mejía, A. and Molina, A. (2007) 'Are we promoting critical autonomous thinking? A discussion on conversational genres and whether they can help us answer this question' in *Cambridge Journal of Education*, 37:3, 409 -426

Merton, R. and Kendall, P (1946) 'The Focused Interview' in *American Journal of Sociology* 51, 541-57

Milot, M. (2005) 'Tolérance, réciprocité et civisme : les exigences des sociétés pluralistes' in F. Ouellet (ed.) *Quelle Formation Pour L'Education á la Religion?* (Saint-Nicolas : Les Presses de l'université Laval)

Moore, G. (1903) *Principia Ethica* (Cambridge: Cambridge University Press)

Morgan, D. L. (1988) *Focus groups as Qualitative Research* (Beverley Hills: Sage Publications)

Munns, G. and Woodward, H. (2006) 'Student engagement and student self-assessment: the REAL framework' in *Assessment in Education* Vol. 13, No.2, 193-213

Narvaez, D. (2005) 'Integrative Ethical Education' in M. Killen and J. Smetana (Eds.), *Handbook of Moral Development* (Mahwah, NJ: Erlbaum) 703-733

Narvaez, D., Bock, T., Endicott, L., and Lies, J. (2004) 'Minnesota's Community

Voices and Character Education Project' in *Journal of Research in Character Education, 2,* 89-112.

Natriello, G. (1987) 'The impact of evaluation processes on students', in *Educational Psychologist*, 22, 155-175

Nesbitt, E. (1991) *My Dad's Hindu, My Mum's Side are Sikhs: Issues in Religious Identity* (Charlbury: National Foundation for Arts Education)

Nesbitt, E. (1998) 'Bridging the Gap between Young People's Experience of their Religious Traditions at Home and School: the Contribution of Ethnographic Research', in *British Journal of Religious Education* 20:2, 102-114

Nesbitt, E. (2004) *Intercultural education: Ethnographic and religious approaches* (Brighton: Sussex Academic Press)

Nesbitt, E. and Jackson, R. (1995) 'Sikh Children's use of "God": ethnographic fieldwork and religious education' in *British Journal of Religious Education* 17:2, 108-120

Nicholson, P. (1985) 'Tolerance as a moral ideal' in J. Horton and S. Mendus (eds.) *Aspects of toleration: philosophical studies* (London: Methuen)

O'Grady, K. (2003) 'Motivation in Religious Education: A Collaborative Investigation with Year Eight Pupils' in *British Journal of Religious Education* Vol. 25:3, 214-225

O'Grady, K. (2005) 'Pedagogy, Dialogue and Truth: Intercultural Education in the

Religious Education Classroom' in R. Jackson and U. McKenna (eds.) *Intercultural Education and Religious Plurality* (Oslo: The Oslo Coalition on Freedom of Religion and Belief), 25-34

O'Grady, K. (2006a) 'The development of Beliefs and Values in Early Adolescence: a case study through Religious Education Pedagogy' in *International Journal of Children's Spirituality* Vol. 11:3, 315-327

O'Grady, K. (2006b) *Motivation in Secondary Religious Education,* University of Warwick, unpublished PhD thesis

O'Grady, K. (2007) 'The Warwick REDCo Community of Practice' unpublished paper, presented at Westhill Seminar, February 2007.

Oancea, A. and Furlong, J. (2007) 'Expressions of excellence and the assessment of applied and practice-based research' in *Research Papers in Education*, 22:2, 119 - 137

OFSTED (2003) *Good Assessment Practice in Religious Education* (London: OFSTED Publications Centre)

OFSTED (2005) *The Annual Report of Her Majesty's Chief Inspector of Schools 2004/2005* at live.ofsted.gov.uk/publications/annualreport0405/4.2.14.html [downloaded 29/11/2006]

OFSTED (2007) *Making sense of religion: A report on religious education in schools and the impact of locally agreed syllabuses* (London: OFSTED)

Organisation for Economic Co-operation and Development (ed.) (2005) *Formative Assessment: Improving Learning in Secondary Classrooms* (Paris: OECD)

Organisation for Security and Co-operation in Europe (2007) *Toledo Guiding Principles on Teaching about Religions and Beliefs in Public Schools* (Warsaw: Organization for Security and Co-operation in Europe)

Ouellet, F. (ed.) (2005) *Quelle Formation pour l'Éducation à la Religion?* (Quebec: Les Presses de l'Université Laval)

Oxfordshire County Council (1992) *Religious Education in Oxfordshire: the Agreed Syllabus* (Oxford: Oxfordshire County Council)

Palmer, G. (1993) 'Politics and Religious education' in *REsource* 16:1 2-6

Parkison, P. (2008) 'Space for performing teacher identity: through the lens of Kafka and Hegel' in *Teachers and Teaching* 14:1, 51-60

Pearce, S. (2004) 'The Development of one Teacher's Understanding of Practitioner Research in a Multi-Ethnic Primary School' in *Educational Action Research* Vol. 12:1, 97-109

Perrenoud, P. (1988) 'La part d'évaluation formative dans toute évaluation continue' in INRAP (ed.) *Évaluer l'évaluation* (Dijon: INRAP) 203-210. Also at http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_1988/1988_05.html [downloaded 10/03/08]

Perrenoud, P. (1991) 'Towards a pragmatic approach to formative evaluation', in Weston, P. (ed.) *Assessment of Pupils' Achievement: Motivation and School Success* (Amsterdam: Swets and Zeitlinger), 79-101

Perrenoud, P. (1998) 'From formative evaluation to controlled regulation of learning processes: towards a wider conceptual field' in *Assessment in Education* Vol. 5:1 85-102

Perrenoud, P. (2005) 'L'évaluation des élèves, outil de pilotage ou pare angoisse?' in *Cahiers pédagogiques,* n° 438, December 2005, 14-16 ; also at www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_2005/2005_10.html [downloaded24/03/2008]

Pike, M. (2002) 'Aesthetic distance and the spiritual journey: educating for morally and spiritually significant events across the art and literature curriculum' in *International Journal of Children's Spirituality* 7:1, 9–21

Pike, M. (2007) 'Values and visibility: the implementation and assessment of citizenship education in schools' in *Educational Review* 59:2, 215-229

Pillonel, M. and Rouiller, J. (2001) 'L'auto-évaluation: une pratique prometteuse mais paradoxale' *Educateur, No. Spécial* 15, 30-32, also at http://www.unige.ch/fapse/SSE/groups/life/textes/Pillonel_Rouiller_A2001_01.html [downloaded 28/02/2007]

Pillonel, M. and Rouiller, J. (2002) 'Faire appel à l'auto-évaluation pour développer l'autonomie de l'apprenant' *Résonances* 7, 28-31, also at

http://www.unige.ch/fapse/SSE/groups/life/textes/Pillonel_Rouiller_A2002_01.html [downloaded 28/02/2007]

Pring, R. (2000) *Philosophy of Educational Research* (London and New York: Continuum)

Pryor, J. and Lubisi, C. (2002) 'Reconceptualising educational assessment in South Africa - testing times for teachers' in *International Journal of Educational Development*, 22, 673– 686

Qualifications and Curriculum Authority (1998) *Exemplification of standards in religious education* (London: QCA)

Qualifications and Curriculum Authority (2000) *Religious Education: Non Statutory Guidance on RE* (London: QCA)

Qualifications and Curriculum Authority (2003) *A Non-Statutory Framework for Religious Education: Report of a Feasibility Study* (London: QCA)

Qualifications and Curriculum Authority (2004) *Religious Education: the non-statutory national framework* (London: QCA)

Radnor, H. (1994) 'The problems of facilitating qualitative formative assessment in pupils' in *British Journal of Educational Psychology* 64, 145-160

Ramprasad, A. (1983) 'On the definition of feedback' in *Behavioural Science* 28:1, 4-13

Religious Education Council (1991) *RE, Attainment and National Curriculum* (Lancaster: Religious Education Council)

RE Online and Culham Institute (2006) *Non-statutory National Framework for Religious Education, 2004 – 'I Can' Attainment Statements,* at http://betterre.reonline.org.uk/assessment/cando.php [Downloaded 26/10/2007]

Reinharz, S. (1997) 'Who am I? The Need for a Variety of Selves in the Field' in R. Hertz (ed.) *Reflexivity and Voice* (Thousand Oaks: Sage Publications), 3-14

Research Assessment Exercise (2006) 'RAE 2008 Panel criteria and working methods: UOA 45, Education' at www.rae.ac.uk/pubs/2006/01/docs/k45.pdf [Downloaded 24/03/2008]

Riding, R. and Butterfield. S (eds.) (1990) *Assessment and Examination in the Secondary School* (London: Routledge)

Robson, G. (1996) 'Religious Education, Government Policy and Professional Practice, 1985-1995' in *British Journal of Religious Education* 19:1, 13-23

Roos, B. and Hamilton, D (2005) 'Formative assessment: a cybernetic viewpoint' in *Assessment in education* 12:1, 7-20

Rowe, M. (1974) 'Wait time and rewards as instructional variables, their influence on language, logic and fate control: Part 1-*Wait Time*' in *Journal of Research in Science teaching*, 11:2, 81-94

Rubin, H. and Rubin, I. (1995) *Qualitative Interviewing: the Art of Hearing Data* (Thousand Oaks: Sage)

Rudge, J. (2000) 'The Westhill Project: Religious Education as Maturing Pupils' Patterns of Belief and Behaviour' in in Grimmitt, M (ed.) *Pedagogies of Religious Education* (Great Wakering: McCrimmonds), 88-111

Sadler, R. (1989) 'Formative Assessment and the Design of Instructional Systems' in *Instructional Science* 18:119-144

Said, E. (1978) *Orientalism* (London: Penguin)

Salmons, P. (2003) 'Teaching or Preaching? The holocaust and intercultural education in the UK' in *Intercultural Education* 14:2, 139-149

Schön, D. (1983) *The Reflective Practitioner* (London: Temple Smith)

Schools Curriculum and Assessment Authority (1994) *Model Syllabuses for Religious Education* (London: SCAA)

Sebatane, E. (1998) 'Assessment and classroom learning: a response to Black and Wiliam' in *Assessment in Education* 5:1, 123-130

Sikes, P. (2000) '"Truth" and "Lies" Revisited' in *British Educational Research Journal* 17:3, 157-167

Sikes, P. and Everington, J. (2000) 'Becoming an RE teacher: a Life History Approach' in *British Journal of Religious Education* 24:1, 8-19

Silverman, D. (1993) *Interpreting Qualitative Data* (London: Sage)

Smith, E. and Gorard, S. (2005) '"They don't give us our marks": the role of formative feedback in student progress' in *Assessment in education* 12:1 pp21-38

Smith, H. and Higgins, S. (2006) 'Opening classroom interaction: the importance of feedback' in *Cambridge Journal of Education* 36:4, 485-502

Smith, L. and Kay, D. (2000) 'Religious Terms and Attitudes in the Classroom (Part 2)' in *British Journal of Religious Education* 22:3, 181-191

Sosa, E. (1980) 'The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge' in *Midwest studies in Philosophy* V: 3-15

Sosa, E. (2003) 'The Place of Truth in Epistemology' in M. DePaul and L. Zagzebski (eds.) *Intellectual Virtue: Perspectives from Ethics and Epistemology* (Oxford: Oxford University Press), 155-179

South Gloucestershire County Council (2000) *Mystery and Meaning* (Gloucester: South Gloucestershire County Council)

Stenhouse, L. (1975) *An Introduction to Curriculum Research and Development* (London: Heinemann)

Stephani, L. (1994) 'Peer, self and tutor assessment: relative reliabilities' in *Studies in Higher Education* 19, 69-75

Stern, J (2006) *Teaching Religious education* (London: Continuum)

Stiggins, R., Griswold, M. and Wikelund, K. (1989) 'Measuring Thinking Skills Through Classroom Assessment' in *Journal of Educational Measurement* 26: 233-246

Sutton, R. (1994) *Assessment for Learning* (Salford: RS Publications)

Swift, N. and Gooding, T. (1983) 'Interaction of wait time feedback and questioning instruction on middle school science teaching' in *Journal of Research in Science Teaching* 20:8, 721 - 730

Thatcher, M. (1993) *The Downing Street Years* (London;HarperCollins)

Tierney, R. (2006) 'Changing practices: influences on classroom assessment' in *Assessment in Education* 13:3, 239-264

Tirri, K. and Nokelainen, P. (2007) 'Comparison of Academically Average and Gifted Students' Self-Rated Ethical Sensitivity' in *Educational Research and Evaluation*, 13:6, 587-601

Torrance, H. (1997) 'Assessment, Accountability and Standards: Using Assessment to Control the reform of Schooling' in A. Halsey, H. Lauder, P. Brown, A. Stuart Wells (eds.) *Education: Culture, Economy, Society* (Oxford; Oxford University Press), 321-331

Torrance, H. and Pryor, J. (1998) *Investigating Formative Assessment; Teaching,*

Learning and Assessment in the Classroom (Buckingham: OUP)

Torrance, H. and Pryor, J. (2001) 'Developing formative assessment in the classroom: using action research to explore and modify theory' in *British Educational Research Journal*, 27, 615– 633

Vygotsky, L. (1987) 'Thinking and Speech' [1934] in *The Collected Works of L.S.Vygotsky, Volume 1, Problems of general Psychology*, ed. R. Rieber and A Carton (New York and London: Plenum Press)

Wakefield Local Education Authority (2001) *Using Pupils' Work to Exemplify Standards of Attainment in Religious Education* (Wakefield: City of Wakefield Metropolitan District Council) (www.wakefield.gov.uk) [downloaded 31/01/2004]

Watson, J. (2007) 'Can children and young people learn from atheism for spiritual development? A response to the National Framework for Religious Education' in *British Journal of Religious Education* 30:1, 49-58

Watts, M. and Ebbutt, D. (1987) 'More than the sum of the parts: research methods in group interviews' in *British Educational Research Journal* 13:1, 25-34

Weber, Max (1978) 'The Concept of "Following a Rule"' in *Selections in Translation* [ed. W. Runciman, trans. E.Matthews] (Cambridge: Cambridge University Press), 99-110

Weiner, G., Arnot, M. and David, M. (1997) 'Is the future female? Female success, Male disadvantage, and changing gender patterns in education' in A.

Halsey, H. Lauder, P. Brown, and A. Stuart Wells (eds.) *Education: Culture, Economy, Society* (Oxford; Oxford University Press), 620-629

Weisse, W. (2007) 'The European Research project on Religion and education "REDCo"' in R. Jackson, S. Miedema, W. Weisse and J-P. Willaime (eds.) (2007) *Religion and Education in Europe: Developments, Contexts and Debates* (Münster: Waxmann)

Wertsch, J. (1985) *Vygotsky and the Social Formation of Mind* (Cambridge and London: Harvard University Press)

Wenger, E. (1998) *Communities of Practice*, (Cambridge: Cambridge University Press).

Westhill College (1989) *Attainment in RE: a Handbook for teachers* (Birmingham: Westhill College)

Westhill College (1991) *Assessing, Recording and Reporting R.E. A Handbook for Teachers* (Birmingham: Westhill College)

Wiliam, D. (1994) 'Towards a philosophy for educational assessment' Paper presented at British educational Research Association conference, Oxford 1994, from www.kcl.ac.uk [downloaded 15/06/2003]

Wiliam, D. (1998) *Enculturating learners into communities of practice: raising achievement through classroom practice*, Paper presented at the European Conference on Educational Research, Ljubljana, from www.kcl.ac.uk [downloaded 15/06/2003]

Wiliam, D., Lee, C., Harrison, C. and Black, P. (2004) 'Teachers developing assessment for learning: impact on student achievement' in *Assessment in Education: Principles, Policy and Practice* 11(1), 49-65

Willaime, J.-P. (2007) 'Teaching Religious Issues in French public schools: from Abstentionist *laïcité* to a return of Religion to Public education' in R. Jackson, S. Miedema, W. Weisse and J-P. Williame (eds.). *Religious and education in Europe: developments, contexts and debates* (Münster: Waxmann), 87-101

Willis, P. (1977) *Learning to labour* (Farnborough: Saxon House)

Winkler, G. (2003) 'Are We Nice, or are We Real? Ethical Issues Emerging from Collaborative Narrative Research' in *Educational Action Research* 11:3, 389-402

Wintersgill, B. (1994) 'The Case of the Missing Models: Exploding the Myths' in RESource 18:1, 6-11

Wintersgill, B. (2000) 'Task-setting in religious education at Key Stage 3: a comparison with History and English' in RESource 22:3, 18-22

Wintersgill, B. (2007) Personal communication

Wittgenstein, L. (1958) *Philosophical Investigations* (Oxford: Oxford University Press)

Wolyie Hussein, J. (2006) 'Hopes and challenges in using action research: the outcome of attempting to help in-service teachers learn how to design, evaluate

ands use reading comprehension questions collaboratively' in *Educational Action Research* 14:3, 377-393

Wright, A. (1998) *Spiritual Pedagogy. A survey, critique and reconstruction of Contemporary Spiritual Education in England and Wales* (Abingdon: Culham College Institute)

Wright, A. (2000) 'The Spiritual Education project: Cultivating Spiritual and Religious Literacy Through a Critical Pedagogy of Religious Education' in M. Grimmitt (ed.) *Pedagogies of Religious Education* (Great Wakering: McCrimmonds), 170-187

Wright, A. (2003) 'The contours of critical religious education: knowledge, wisdom, truth' in *British Journal of Religious Education*, 25 (4): 279-91

Wright, A. (2004) *Religion, Education and Post-modernity* (London: RoutledgeFalmer)

Wright, A. (2008) 'Contextual religious education and the actuality of religions' in *British Journal of Religious Education* 30:1, 3 - 12

Zagzebski, L. (1996) *Virtues of the Mind* (Cambridge: Cambridge University Press)

Zagzebski, L. (2003) 'Intellectual Motivation and the Good of Truth' in M. DePaul and L. Zagzebski (eds.) *Intellectual Virtue: Perspectives from Ethics and Epistemology* (Oxford: Oxford University Press)

Østberg, S. (2000) 'Islamic Nurture and identity management: the Lifeworld of Pakistani Children in Norway' in *British Journal of Religious Education* Vol. 22:2, 91-102

Østberg, S. (2003) 'Cultural diversity and common citizenship: reflections on ethnicity, religion, nationhood and citizenship among Pakistani young people in Europe' in R. Jackson (ed.) *International perspectives on Citizenship, Education and Religious Diversity* (London: RoutledgeFalmer)