

MATHEMATICS
TEACHERS' CLASSROOM
ASSESSMENT PRACTICES:
A CASE STUDY IN A
MALTESE
SIXTH FORM COLLEGE

MICHAEL A. BUHAGIAR

Thesis submitted to the Faculty of Education,
University of Nottingham, for the degree of
Doctor of Philosophy

December 2005

DEDICATION

To all my loved ones,
especially mum and dad,
for their continued support and understanding

ACKNOWLEDGEMENTS

My sincere thanks go primarily to Professor Roger Murphy, Dr Malcolm Swan and Professor Ronald Sultana for generously offering their expertise, guidance and support all along my seemingly ‘endless’ research journey. I moved ahead only because I could tap on their wealth of knowledge, experience and readiness to help. I am also grateful to Professor Kathleen Hart and Professor Mark Borg for their valid contributions in the early stages of the research.

I am grossly indebted to my participant-colleagues at school, many of who are also my friends, for initially inspiring this research and for subsequently contributing generously towards its development. My thanks are also due to Noel Buttigieg for kindly accepting to participate in the piloting phase of the study.

My visits to Nottingham would not have been half as rewarding without the friendship and support that I received from a number of fellow students. In particular, I treasure my friendships with Hanan Innabi, Siew Ming Tang, Satomi Iwasawa, Alan Chow, Jenny Setiawan, Aysun Yavuz, Carole Mallia and Cathy Gibbons. I also discovered in Nottingham how good it feels to socialise with Maltese – namely, Adrian and Jackie, Marco and Simone (and later baby Andrew), and Cedric and Kathleen – away from Malta.

I am very thankful to Wilf Hastings – a true gentleman who has now sadly passed away – for kindly welcoming me at his home in Beeston on my last two visits to Nottingham.

My sincerest thanks go to Deborah Chetcuti for being a true friend and a constant guiding light along my journey. I am also very grateful to Grace Grima and Mauro Scerri for generously offering their help without me even having to ask. Thanks also to Leonard Bezzina and Cettina Axiak for providing some of the reading materials.

Finally, I would like to thank the University of Malta for partially funding my studies.

TABLE OF CONTENTS

Title Page	i
Dedication	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables and Figures	xi
Abstract	xii
Glossary	xiii
Acronyms	xv
An Invitation to a Research Journey	1
PART ONE: Setting the Scene	3
Chapter 1: Introducing the Research	4
1.0 Presenting Myself	4
1.1 The Early Stages of the Research	4
1.1.1 Deciding to Read for a PhD	4
1.1.2 Undergoing a Personal Paradigm Shift	5
1.2 The Research Area	8
1.2.1 Selecting the Research Area	8
1.2.2 Identifying the Research Area	9
1.3 The Research Questions	10
1.3.1 Presenting the Research Questions	10
1.3.2 Broadening out the Underlying Issues	11
1.4 The Structure of the Thesis	12
Chapter 2: The Research Background	15
2.0 An Introduction to the Maltese Islands	15
2.1 The Framework and Structure of the Maltese Educational System	15
2.1.1 The Framework	15
2.1.2 The Structure	16
2.2 Certification and Assessment in Malta	17
2.2.1 Certification	17
2.2.1.1 At SEC Level	18
2.2.1.2 At Matriculation Level	18
2.2.2 Assessment	19
2.2.2.1 Assessment that is Traditionally Based on Examinations	19
2.2.2.2 Growing Uneasiness with Malta's Examination Culture	19

2.3	The Intermediate Level of the Matriculation Certificate	21
2.3.1	‘Defining’ the Intermediate Level	21
2.3.2	The Pure Mathematics at Intermediate Level (PMI) Option	21
2.4	The Junior College of the University of Malta	22
2.4.1	Rationale, Aim and Structures	22
2.4.2	Entry Requirements and its Impact on the Educational System	23
2.4.3	The Academic Calendar, Regulations and Teaching Methods	23
2.4.4	The Evaluation and Promotion System	24
2.5	The Department of Mathematics at the Junior College	25
2.5.1	The Physical Environment	25
2.5.2	Teachers and Operational Dynamics	26
2.5.3	The PMI Option	27
2.6	Rationale for Change in the Maltese Educational System	28
2.6.1	The Current Situation	28
2.6.2	The Vision Ahead	30
2.6.3	Positioning the Thesis	31
	PART TWO: The Literature Review	32
	Introduction: Learning and Assessment	33
	Chapter 3: Teachers and Classroom Assessment	41
3.0	The Need to Move Away from the Traditional Assessment Paradigm	41
3.1	The Alternative Assessment Paradigm	45
3.1.1	Philosophical Underpinnings of the Alternative Assessment Paradigm	45
3.1.2	Mathematics and the Alternative Assessment Paradigm	48
3.2	Classroom Assessment	50
3.2.1	Defining Classroom Assessment	51
3.2.2	The Purposes of Classroom Assessment	53
3.2.3	Classroom Assessment in the Alternative Assessment Paradigm	56
3.2.4	Assessment for Learning inside the Classroom: Where do we Stand?	61
3.3	Teachers as Classroom Assessors	62
3.3.1	The World of Teachers	63
3.3.2	Teacher Competence in Classroom Assessment	66

3.3.3	Models of Classroom Assessment by Teachers	70
3.3.3.1	The Torrance and Pryor (1998) Study	70
3.3.3.2	The Gipps et al. (1995) Study	71
3.3.3.3	What do these Studies tell us about the Assessment Reform Process?	73
3.3.4	Classroom Assessment by Teachers in relation to 'Others'	75
3.3.5	Classroom Assessment by Teachers in relation to School and External Assessment	78
3.4	The Classroom Assessment Cycle within the Alternative Assessment Paradigm	81
3.4.1	Planning the Activity	82
3.4.2	Gathering Evidence	83
3.4.2.1	Observation	84
3.4.2.2	Communication	85
3.4.2.3	Tasks	88
3.4.2.4	Class Tests	90
3.4.3	Interpreting Evidence	92
3.4.4	Using Evidence	93
3.4.4.1	Instructional Decisions	93
3.4.4.2	Feedback	94
3.4.4.3	Grading	96
3.5	Looking Forward with Hindsight	98
	PART THREE: The Methodology	99
	Chapter 4: Exploring the Issues	100
4.0	Dealing with Subjectivity	100
4.1	Reflections on Ontology, Epistemology and Methodology	100
4.1.1	Ontological Reflections	101
4.1.2	Epistemological Reflections	102
4.1.3	Methodological Reflections	103
4.2	Selecting the Research Design and Methods	105
4.2.1	Opting for a Qualitative Case Study	105
4.2.2	Using Multiple-Methods within a Flexible Research Framework	106
4.3	Presenting the Data Collection Methods	108
4.3.1	Documents	108
4.3.2	Observations	110
4.3.3	Interviews	112
4.3.3.1	Deciding on the Interview Format	113
4.3.3.2	Qualitative Interviewing	115
4.4	Ethical Considerations	116

Chapter 5: Implementing the Study	120
5.0	The Research Site 120
5.1	The Pre-Fieldwork Phase 121
5.1.1	Entering the Field 121
5.1.2	Ethical and Pragmatic Rethinking 122
5.2	Formally Entering the Field 123
5.2.1	Gaining ‘Formal’ Access 123
5.2.2	Selecting Participants and Building Trust 125
5.2.3	Contacting the Participants 126
5.2.4	Seeking Informed Consent 127
5.2.5	Renegotiating Access with Participants 129
5.3	The Fieldwork Phase 130
5.3.1	The Fieldwork Journal 130
5.3.2	The Interviews 131
5.3.2.1	Preparing the Interviews 131
5.3.2.2	The Purposes of the Interviews 132
5.3.2.3	Conducting the Interviews 133
5.3.2.4	Transcribing the Interviews 134
5.3.3	The Classroom Observations 135
5.3.4	Leaving the Field 136
5.3.5	Reflections on my Insider Status 136
5.4	Turning Data into Research Findings 137
5.4.1	Analysing and Interpreting the Data 137
5.4.2	Key to Interpreting the Research Findings 139
PART FOUR: The Data Analysis	140
Chapter 6: Meeting Teachers and Entering Classrooms	141
6.0	Deciding to Become One of the Participants 141
6.1	Introducing the Teachers 142
6.2	The Rationale behind Michael’s Story 145
6.3	Getting to Know the Teachers 147
6.3.1	Their Views on Teacher Education 147
6.3.2	Their Views on Teaching Experience 149
6.3.3	Their Views on the School 150
6.3.4	Their Views on the Mathematics Department 151
6.3.5	Their Views on the PMI Option 153
6.4	Classroom Action, Interaction and Teaching Priorities 156
6.5	The Typical PMI Lesson 159
6.5.1	The Teachers speak of ‘Teaching’, not ‘Lecturing’ 159
6.5.2	The Two PMI Teaching Systems 160
6.5.3	The PMI Lesson – a Cycle of Exposition, Practice and Consolidation 160

Chapter 7: Teachers' Classroom Assessment Practices	163
7.0	The Origins of TCAP 163
7.1	Roles and Relationships in TCAP 164
7.1.1	The Role of the Classroom Teacher 164
7.1.2	TCAP and 'Others' 165
	7.1.2.1 TCAP and Colleagues 166
	7.1.2.2 TCAP and Students 168
7.2	Ways in which Teachers obtain Information 171
7.2.1	Collating Information 171
7.2.2	Generating Information 172
	7.2.2.1 Observing Students 173
	7.2.2.2 Checking Students' Work 175
	7.2.2.3 Listening to Students 177
	7.2.2.4 Asking Questions 179
	7.2.2.5 Class Tests 181
7.3	Formal and Informal Sources of Information 184
7.3.1	Information gained from Formal and Informal Sources 184
7.3.2	Complementarity and Agreement between Formal and Informal Sources 186
7.3.3	The Use of Formal and Informal Sources for Official Reporting 188
7.4	The Classroom Assessment Tasks 191
7.4.1	The Centrality of Tasks in Classroom Assessment 191
7.4.2	The Criteria for Selecting Classroom Tasks 192
7.4.3	The Nature of Classroom Tasks 195
7.5	The Results of Classroom Assessment 197
7.5.1	Recording the Results 197
7.5.2	Responding to the Results 199
7.5.3	Communicating the Results 203
7.6	The Beneficiaries and Benefits of Classroom Assessment 205
7.6.1	The Benefits to Teachers 206
7.6.2	The Benefits to Students 210
7.7	Constraints and Explanations 214
7.7.1	Teachers on the Quality of their Assessment Practices 214
7.7.2	The Constraints on TCAP 216
7.7.3	Explanations for TCAP 218
7.7.4	TCAP in a Constraint Free Environment 220
PART FIVE: Discussion and Conclusions	223
Chapter 8: Exploring the Implications of TCAP for Learning	224
8.0	Getting Focussed 224
8.1	TCAP – Revisiting the Main Findings 224
8.1.1	Assessment that Lacks an Effective Educational Base 224
8.1.2	Assessment that Pursues Traditional Forms and Approaches to Teaching 225
8.1.3	Assessment that Emphasises Non-Professional Functions 226

8.1.4	Assessment that is Firmly in Teachers' Hands	228
8.1.5	Assessment that is Characterised by Collegial Isolation	229
8.1.6	Assessment that is Not Well Integrated with Teaching and Learning	230
8.1.7	Assessment that Provides Teachers with Surface Knowledge of Students	231
8.1.8	Assessment that Provides Students and Outsiders with Limited Feedback	233
8.2	The Impact of Teacher Beliefs on TCAP	234
8.2.1	Teacher Beliefs about the Educational Process and TCAP	235
8.2.2	Teacher Beliefs about Mathematics and TCAP	236
8.3	The Implications of TCAP for Learning	238
8.3.1	When Assessment Lacks an Effective Educational Base	238
8.3.2	When Assessment Pursues Traditional Forms and Approaches to Teaching	239
8.3.3	When Assessment Emphasises Non-Professional Functions	240
8.3.4	When Assessment is Firmly in Teachers' Hands	241
8.3.5	When Assessment is Characterised by Collegial Isolation	242
8.3.6	When Assessment is Not Well Integrated with Teaching and Learning	243
8.3.7	When Assessment Leads to Surface Knowledge of Students	244
8.3.8	When Assessment Provides Limited Feedback	245
8.4	The Quality of Teachers' Classroom Assessment Practices	246
8.4.1	What the Teachers Say	247
8.4.2	What the Research Findings Suggest	248
8.5	Recognising the Need for Change	249
Chapter 9: Moving towards Assessment for Learning		251
9.0	The Reform Process in Malta: What do the Research Findings Suggest?	251
9.1	The Change Process	252
9.1.1	Recognising that Change is a Complex Process	252
9.1.2	Recognising that Change is Possible	253
9.2	Looking beyond Teachers' Explanations	255
9.2.1	Revisiting Teachers' Explanations	255
9.2.2	A Critical Look at Teachers' Explanations	255
9.2.3	Identifying Three Levels of Concern	257
9.3	Acknowledging the Difficulties	258
9.3.1	At Teacher Level	258
9.3.2	At School Level	260
9.3.3	At National Level	262

9.4	Exploring Prospects for Improvement	264
9.4.1	The What of Change	264
9.4.2	The How of Change	265
9.4.2.1	At Teacher Level	266
9.4.2.2	At School Level	267
9.4.2.3	At National Level	269
9.5	Concluding Comments	271
	My Journey as a Teacher Continues ...	273
	References	276
	Appendix I: The Contrasting Pairs of Statements and Cards	296
	Appendix II: The Four Tasks	298

LIST OF TABLES AND FIGURES

TABLES

Table 2.1	Teacher allocation inside offices	26
Table 3.1	NCTM assessment standards for school mathematics	49
Table 3.2	The ten ‘assessment for learning’ principles	57
Table 6.1	Student classroom participation that is valued by teachers	157
Table 7.1	Forms of limited assessment interaction amongst teachers	167
Table 7.2	Number of class tests during Year 2 of the study	182
Table 7.3	A class test ‘experiment’	183
Table 7.4	Teachers’ reasons as to why test marks may be misleading	186
Table 7.5	Teachers’ comments on the four tasks	193
Table 7.6	The task selection criteria by the three determining contexts	194
Table 7.7	The nature of classroom tasks	196
Table 7.8	Teachers’ work of consolidation	200
Table 7.9	The correlation coefficients based on teachers’ predictions	208
Table 7.10	The potential benefits of class tests for students	213
Table 7.11	Constraints on classroom assessment	217
Table 7.12	The ‘new’ TCAP in a constraint free environment	221

FIGURES

Figure 2.1	The layout of the mathematics department	25
Figure 3.1	The four phases of classroom assessment	82

ABSTRACT

This study explores the classroom assessment practices of twelve teachers teaching Pure Mathematics at Intermediate Level in a Maltese sixth form college. Apart from identifying the participants' assessment views and practices (including how they explain their practices), the focus is on examining the implications that these views and practices have for learning. In order to gain deeper insights into the underlying issues, the researcher – who is also a colleague of the participants – assumed the additional role of 'thirteenth participant'. This duality proved particularly useful in understanding how the links between classroom assessment and learning can be strengthened.

Using a methodological approach that draws on ethnography, the data was collected over a three-year fieldwork period through documents, observations and semistructured interviews. In the course of this reflective and reflexive research journey, characterised throughout by an understanding of the subjectivity of meanings, every effort was made to observe the canons of qualitative research. The analysis of the data, which began germinating in the very entries of the fieldwork journal, was carried out by themes. Initial efforts aimed at submitting the emerging themes from the research to the participants in order to check on the interpretations made and to generate further insights had to be discontinued.

The research evidence suggests that, in spite of the chorus of voices and policies clamouring for the strengthening of the formative dimension of classroom assessment, the classroom assessment practices of the teachers studied are still very much grounded in the traditional assessment paradigm. It is subsequently argued that this continuing emphasis on assessment as a form of measurement for managerial and accountability purposes weakens the learning aspect – which should after all be the focal point of all educational activities – inside classrooms. In the search as to how classroom assessment can become more supportive of learning, three levels of difficulties – namely, the teacher, school and national contexts – were identified. The study finally suggests how the principles that underline the emerging assessment paradigm can start gaining roots within the complex realities that teachers face.

GLOSSARY

Assessment Marks: At the end of each term, teachers at the Junior College forward an assessment of each student, in the form of a grade, to the school administration. These Assessment Marks must be out of 100 and in multiples of 5.

College Board: This board, which governs the Junior College, is responsible for: (i) the attainment of the Junior College aims; (ii) the monitoring of the general and financial administration of the Junior College; (iii) the academic appointments at the Junior College; and (iv) collaboration with the MATSEC Board, the Education Division and analogous bodies.

Education Division: It represents the administrative arm of the government in educational matters concerning curriculum, examinations, planning and development, operations, finance and administration, further studies and adult education, and student services.

End-of-first-year examination: As part of the Junior College promotion system, first year students sit for end-of-year examinations that carry a weighting of 70 %. The remaining 30% are based on Assessment Marks.

Faculty of Education: The setting up of the faculty in 1978 within the University of Malta effectively raised local teacher education from certificate to degree level.

Intermediate Level: This level was introduced in 1995 by MATSEC when it launched the Matriculation Certificate. It supposedly represents one-third of an Advanced Level.

Junior College: The sixth form college of the University of Malta that was set up in 1995 to prepare prospective university students over a two-year study programme.

Maintenance grant: All students who follow full-time courses in any post-secondary institution or the University of Malta are entitled to a maintenance grant. These grants are driven by the philosophy that ‘investing in today’s youngsters is a guarantee for the country’s future’.

Matriculation and Secondary Certificate Examinations Board (MATSEC): The examination body that organises the 16+ Secondary Education Certificate and the 18+ Matriculation Certificate examinations under the umbrella of the University of Malta

Matriculation Certificate: It is MATSEC’s adaptation of the International Baccalaureate (IB). Matriculating students are required to study two subjects at Advanced Level and four subjects at Intermediate Level. The Certificate’s six subjects are assessed in one examination session. Holders of the Matriculation Certificate are eligible to join undergraduate courses at the University of Malta.

National Minimum Curriculum (NMC): The legal document that describes the aims of education, the subject content to be taught and the methods in which it should be taught.

Paper 2B: This Secondary Education Certificate (SEC) examination paper – which includes less demanding questions than those set in Paper 1, and even less demanding questions than those set in Paper 2A – was introduced by MATSEC as part of its differentiated papers strategy. Candidates sitting for Paper 1 and Paper 2A may qualify for grades 1 to 5, and candidates sitting for Paper 1 and Paper 2B may qualify for grades 4 to 7.

Personal contact hour: Junior College students have the possibility to meet and discuss difficulties with their teachers on a one-to-one basis during personal contact hours.

Principal: The principal is appointed by the University Council from amongst the subject coordinators to see to the day-to-day running of the Junior College.

Registrar: The appointed person is in charge of the non-academic staff offering support services at the Junior College.

Secondary Education Certificate (SEC): It is awarded to students who pass MATSEC examinations at the end of their secondary schooling. SEC examinations are modelled on the General Certificate of Secondary Education (GCSE) examinations in the United Kingdom.

Subject coordinator: Following consultations with the teaching staff, the Junior College administration appoints a subject coordinator within each department responsible for its day-to-day administration.

Teacher's Certificate: Until the early 1970s, two residential Teacher Training Colleges – one for men and another for women – used to offer a two-year course leading to a Teacher's Certificate.

University of Malta: This university – the only one on the island – traces its origins to a Jesuits' college founded in 1592. It has at present circa 8000 enrolled students, including some 500 foreigners, following full-time or part-time degree and diploma courses organised by its ten faculties and numerous institutes.

ACRONYMS

ARG	Assessment Reform Group
BEd	Bachelor of Education
BSc	Bachelor of Science
EU	European Union
GCSE	General Certificate of Secondary Education
IB	International Baccalaureate
MATSEC	Matriculation and Secondary Education Certificate Examinations Board
MEd	Master of Education
MSc	Master of Science
MSEB	Mathematical Sciences Education Board
MUT	Malta Union of Teachers
NCTM	National Council of Teachers of Mathematics
NMC	National Minimum Curriculum
OFSTED	Office for Standards in Education
PGCE	Post Graduate Certificate in Education
PhD	Doctor of Philosophy
PMI	Pure Mathematics at Intermediate Level
SEC	Secondary Education Certificate
TCAP	Teachers' Classroom Assessment Practices
TGAT	Task Group on Assessment and Testing
ZPD	Zone of Proximal Development

An Invitation to a Research Journey

I joined the Junior College of the University of Malta¹ as a teacher of Pure Mathematics when it was set up in 1995. The opening of this university sixth form college coincided with the local move at pre-university level away from the English three Advanced Levels system towards a wider 16+ curriculum. Although I arrived at the Junior College with a rich and varied baggage of teaching experiences – in the previous ten years I had taught at both primary and secondary levels – I felt distinctly novel, unprepared, and in need of reassurance. As a result, I suddenly lost my self-confidence and most of my free time. Indeed, I had to spend long hours of preparation for each lesson, never too much ahead of my students and occasionally fearing the worst.

I spent my first year at the Junior College too engrossed in lesson preparations to be that bothered with anything else. But later, as I gained in confidence and experience, I started to widen my horizons and take greater interest in what was going on outside my classrooms, in particular within the mathematics department. I was particularly impressed by the ‘interest’ surrounding the Intermediate Level courses that had just been introduced as part of the new, wider sixth form curriculum. Indeed, this new level – that is somewhat in between the traditional Ordinary and Advanced Levels – proved to be an area of constant discussion and agitation amongst my colleagues, and I could sense an air of rejection about it (see section 2.3.1). This was the spark that led to this research – my explorative curiosity had been aroused. The Intermediate Level has since retained an important position within the study even though the research focus has shifted from my initial intention to evaluate the Pure Mathematics at Intermediate Level (PMI) option, which I also teach, to exploring PMI teachers’ classroom assessment practices.

¹ This open disclosure of the name of the research site arises from my impossibility to shield its identity from anyone with even a basic knowledge of Maltese post-secondary institutions, as even a simple reference to the number of participants involved would immediately identify it. However, as part of the measures to protect individual identities, at least from school outsiders, I have eliminated from the write-up any direct reference to the research dates. At the Junior College, whilst new teachers have joined over the years and others have left, teachers are appointed to administrative positions on a rotational basis. This means that persons within cannot be easily pinpointed without an accurate timeframe. Nevertheless, as the chronology of my study has to be respected and my long permanence in and off the field recorded, I am fielding ‘dummy’ dates instead. Thus, ‘Year 1’ refers to the first scholastic year (i.e., September to August) of the research, ‘Year 2’ to the second, and so on.

What follows is an account of my research journey – written in the first person in recognition of my ‘active’ role (Wolcott, 1990) – that has left a strong, indelible mark upon myself. Not only has this journey, which is basically an insider study that draws on ethnographic methodology, challenged and changed my belief system, but it has also brought me closer to understanding myself. I have found that by exploring those with whom I share my professional life, I also learned about myself. My only guarantee is an honest quest for knowledge based on long hours of work, dedication and respect towards my school, colleagues, the reader and myself. Although I initially set out to explore and reflect upon the views and practices of my colleagues teaching the PMI option, it eventually became clear that my understanding would benefit should I – as I did – also begin to consider myself as one of the participants and act accordingly.

Still, as my resulting knowledge is necessarily partial and contextual, I do not assume in my own story the role of the authoritative and omniscient narrator who tells the whole truth and nothing but the truth (Elbaz, 1990). My account thus “needs to be read as a conscious selection written for a particular purpose, which can be challenged on the grounds of interpretation and meaning rather than on the basis of falsification of a fixed ‘truth’” (Weiner, 1994, p. 11). But whilst my perspectives embedded within it are admittedly partial and thus open to criticism, they have been honestly assumed. They are, like all the work that I do, the products of the interrelationship between my personal biography, my place in the social structure, and the cultural milieu and historical period in which I live (Weiner, 1994).

I can now begin my story ...

Part ONE

SETTING THE SCENE

CHAPTER 1

Introducing the Research

1.0 Presenting Myself

The awareness that all research reflects a partisanship that derives from the social identity and values of the researcher (Troyna, 1995; cited in Bassey, 1999) encourages me to begin the journey by introducing myself to the reader. I am a single Maltese male in my early forties. I grew up in a caring, working class family. Were it not for the insistence and encouragement of my mother – undoubtedly the single most influential person in my life – I would probably not even have bothered to proceed to sixth form education, let alone university. But she kept faith in me throughout, even though I never really excelled as a student at primary and secondary levels. I was then in what teachers usually refer to as the ‘can do better category’. Moreover, there were times when I did not do so well, especially in mathematics, the subject that I was to ‘discover’ later midway through my secondary schooling. Discovering my ability to do very well in mathematics did wonders to my self-esteem and probably paved my way to a rewarding university education (before embarking on my doctoral studies, I graduated BEd in 1985 and MEd in 1991 from the University of Malta) and teaching career (apart from teaching at primary, secondary and post-secondary levels, I have also lectured part-time in pre-service and in-service programmes for mathematics teachers). In the meantime, I have also: (i) participated in research projects; (ii) supervised undergraduate research in mathematics education; (iii) contributed in local popular and professional journals; (iv) collaborated with the teachers of a local secondary school to help review and improve their mathematics programme; and (iv) involved myself in local 16+ mathematics certification examinations.

1.1 The Early Stages of the Research

1.1.1 Deciding to Read for a PhD

Although I had been toying with the idea of reading for a PhD for some time, it was the general ‘commotion’ surrounding the introduction of the Intermediate Level

courses (see section 2.3.1), particularly the PMI option, that eventually catapulted me into action. I soon realised however that my background as a reader of educational literature had still not prepared me for the uneven, stumbling and wavering progress that characterises field research. My experience suggests that although field research, with its emphasis on seeing ‘what’s there’, sounds simple, it is not (Ely et al., 1991). The difficulties I have had to face emerge clearly from what I wrote at one stage in my fieldwork journal when I was feeling particularly down:

After two years at it, I may as well be frank about my ‘ulterior’ motives for doing a PhD. It seemed trendy (many of my friends were doing it); I was bored (I was fed up with watching TV and repeatedly going out to meet the same people in the same places); I thought I could do it (I imagined I had the brains, the time and the money for it); and it seemed like a future investment (it would provide me with alternative career prospects). But I’m afraid I still have to figure out why I’m still doing it! (August, Year 2)

When I wrote this I was seriously considering whether I should discontinue my studies. I felt then that although time was passing by, I was still not making any significant headway. Even my relationship with my supervisors appeared strained at that moment. I felt misunderstood and, at times, I could not really figure out what they were expecting from me. Maybe they were just quietly waiting and hoping for me to find myself. I still had not appreciated how personal doing educational research is in reality. Only later did it sink in that this frustratingly long period of seemingly lack of progress was actually the gestation period I required to move ahead. The ‘sad’ message in my fieldwork journal has long ceased to signify ‘desperation’. Instead, I see it now as signalling ‘hope’ – a reminder that perseverance pays.

1.1.2 Undergoing a Personal Paradigm Shift

A research paradigm conditions the thinking patterns of researchers working within it and underpins their research actions (Bassegy, 1999). For each paradigm carries its own assumptions about the world and existence (ontology), about knowledge and learning (epistemology), and about how knowledge is obtained (methodology) (Ernest, 1998). I realise with hindsight that my inability to make good progress in the first two years of the study can be largely attributed to my not having considered seriously enough these issues beforehand. I had embarked on the project with an unresolved internal

philosophical struggle – a matter that needed to be addressed if I were to register significant progress in the research. I became aware that I could not hope to understand anything that is outside of me without first exploring where I stand on issues in relation to myself as a person and as an educational researcher.

Nowadays, authors (e.g., Bassey, 1999; Creswell, 1994; Ely et al., 1991; Hammersley & Atkinson, 1995; Maykut & Morehouse, 1994; Miles & Huberman, 1994; Silverman, 1993) normally cluster paradigms under two overarching philosophical positions – the positivist and interpretive paradigms.² My main research experiences prior to the present project were the BEd and MEd dissertations – both of which reflect my then positivist philosophical outlook that contends that there is a reality out there to be studied, captured and understood (Denzin & Lincoln, 1994). Believing in the existence of an ‘objective reality’ that exists outside ourselves and is independent of the knower (Lincoln & Guba, 1985), I used to think that ‘reality’ can be approximated, and consequently be better understood, over a number of studies that apply good principles of investigation (read ‘quantitative’). Apart from the ‘cultural heritage’ of positivism (Maykut & Morehouse, 1994), I can think of two main reasons for my previous acceptance of this worldview. First, my overwhelmingly behaviourist-oriented initial teacher-training education at university, and second, my specialisation in statistics at undergraduate level mathematics. Still not realising that even statistics can only result in yet another constructed reality (Lincoln & Guba, 1985), I erred by confusing the complex mathematical formulae and tests for significance with indicators of importance (Maykut & Morehouse, 1994).

The first significant cracks to my positivist outlook on life appeared whilst I was doing the MEd degree. As opposed to my undergraduate experience, here I encountered local people who had actually conducted research within the interpretative paradigm and who spoke enthusiastically, convincingly and competently about it. I grew to appreciate that their work, apart from being well grounded, is a source of rich

² I use the terms ‘positivist’ and ‘interpretive’ as umbrella paradigm concepts. Different authors have referred to these two main philosophical positions by various other names. For instance, the positivist paradigm is also known as conventional, dominant, experimental, traditional and quantitative, and other names for the interpretive paradigm include alternative, constructivist, emerging, hermeneutic, naturalistic, phenomenological and qualitative. In this study, whilst these terms are used interchangeably within each of the two main paradigmatic fields, after Guba and Lincoln (1994) I reserve the terms ‘quantitative’ and ‘qualitative’ to types of methods rather than philosophical positions.

descriptions and explanation of processes in identifiable local contexts (Miles & Huberman, 1994). Moreover, as their work was embedded in everyday constraints, it seemed 'more real' than the sterile and clinical world of experiments to which I had been previously exposed. For the first time, I began to wonder if "our knowledge of ourselves, our thoughts and actions, have the same character as our knowledge of the terrain at our feet and the material world about us" (Hollis, 1994, p. 9). This made me question in turn the validity of using experimental methods, originally developed and employed to study the natural world, to study human consciousness and action.

I was beginning to understand better that as "people's behaviour is not caused in a mechanical way, it is not amenable to the sort of causal analysis and manipulation of variables that are characteristic of the quantitative research inspired by positivism" (Hammersley & Atkinson, 1995, p. 8). But my gradual refusal to continue privileging science as universal, superior and final had provoked within me a feeling of losing profound certainties (Hughes & Sharrock, 1997). Even at the technical level, the profound security that my statistical expertise had previously given me no longer applied. Ironically, I had embarked on my PhD studies reduced to a novice in need of resocialisation from my early and intense exposure to the received view of science (Guba & Lincoln, 1994). At first, I thought that I could get over my uneasiness by adopting a learning-by-doing approach. However, I eventually realised that each time I met a new 'problem' – a frequent enough occurrence when reading for a PhD – I was continually falling back on the traditional research model that I knew best (Maykut & Morehouse, 1994).

At this point, I understood that if I was going to "swim against the tide of cultural heritage of how science is done" (Maykut & Morehouse, 1994, p. 2), I needed to entrench my still embryonic philosophical position on sound theoretical grounds. Thus, I 'withdrew' temporarily from my research proper, and concentrated instead on philosophical and methodological readings and reflections. This was to be the final and decisive phase of my 'interpretive turn' – my epistemological shift away from positivism and towards interpretivism (Howe, 1998). Having adequately dealt with this 'difficulty', my research took on a new lease of life. I was able from then onwards to adopt research approaches and techniques that complement and demonstrate my underlying philosophical preconceptions (Hughes & Sharrock, 1997).

1.2 The Research Area

1.2.1 Selecting the Research Area

When I began this project, my desire was to do something ‘useful’ – what Hammersley (2003) calls ‘educative’ – in the sense of being specifically designed to improve educational policymaking or practice. It also had to be a piece of research that would satisfy me both intellectually and emotionally. In evaluating the newly launched PMI option I saw the opportunity to target these aims. For I could sense that: (i) there was a problem that needed addressing (as everyone was complaining about it); (ii) it was something that interested me at a personal level (as I was a PMI teacher myself); and (iii) my research would prove useful (as all innovations need to be evaluated). Although this was my original intention, using what Hammersley and Atkinson (1995) call the researcher’s ‘exercise of judgement in context’, I ended up doing something quite different in reality. This became possible as I worked within a flexible research framework, viewing the research design as a reflexive process that operates throughout every stage of the project. Within my emergent and integrated research design that draws on ethnographic principles, the data collection, hypotheses construction and theory building were not three separate phases, but were interwoven into one another.

Thus, although I began the study with a particular focus in mind, I could still refine and alter this focus as new information made it relevant to do so (Lincoln & Guba, 1985). In particular, I collected and examined my data with the view that research questions can and should change, and to love that possibility (Ely et al., 1991). My change in focus from evaluating the PMI option to exploring the classroom assessment practices of PMI teachers in reality knows of many small intermediary steps. But with each new step I was getting closer to the captivating world of educational assessment in general, and to teachers’ assessment practices in particular. Step by step, I found myself immersing in an area to which I had never given much thought before. Apart from the emerging data itself, I would include amongst the more important influences behind my change in research direction:

- Being assigned two new supervisors – both keenly interested in assessment – during the first year of the study;

- Interactions with a close friend who was finalising her PhD in educational assessment at the time;
- Changes in Malta's National Minimum Curriculum (NMC) that for the first time brought formative assessment at the centre of local educational discourse;
- My realisation, both as a teacher and as a reader of educational literature, that assessment has much to offer to teaching and learning;
- The PMI teachers' serious preoccupations about the then still 'unknown' external assessment, and their apparent 'disinterest' in classroom assessment matters.

My research focus has evidently varied in response to the emerging field data as well as to personal and contextual changes. Moreover, during this period I also began to appreciate that “the primary goal of research is, and must remain, the production of knowledge” (Hammersley & Atkinson, 1995, p. 17). Thus, my aim now is to produce research that is informative rather than educative. Informative research, in contrast with educative research that aims at changing people in some respect and is specifically designed to do this, aims solely at providing information that could be relevant to people's needs (Hammersley, 2003).

1.2.2 Identifying the Research Area

Assessment, in its various guises, has penetrated into almost every aspect of human endeavour (Broadfoot & Black, 2004). In particular, assessment in education – which is the process of gathering, interpreting, recording and using information about students' responses to an educational task (Harlen et al., 1992) – now enjoys growing centrality within educational discourse due to its widespread and significant role.

... assessment is a powerful tool: it can shape curriculum, teaching and learning; it can affect how pupils come to see themselves both as learners and in a more general sense as competent or not; through labelling and sorting pupils (certification and selection) it affects how pupils are seen by others; it controls access to further education and high status careers. (Gipps, 1994, p. 144)

It may be in recognition of this that, as Broadfoot (1996) points out, assessment issues have grown over the past few years from relative insignificance into one of the most prominent features of many governments' educational strategy. Apart from the recognition that “good education, by definition, encompasses good assessment” (Murphy & Torrance, 1988, p. 7), there is now a substantial and rapidly growing literature output on assessment (Broadfoot, 1996).

Far from being a purely academic or dispassionate scientific process, assessment is a process of interpersonal communication with personal effects (Stiggins & Conklin, 1992). Assessment is not an activity that can be done to students, but is accomplished by means of social interaction in which the practices of the participants have a critical effect on the outcomes of assessment, which are thus actively produced rather than revealed and displayed by the assessment process (Pryor & Torrance, 2000). For certain, the stakes involved are high and can carry long-lasting consequences, especially for the students. So much so that assessment defines “the attitude students take towards their work, their sense of ownership and control of their own learning: the strategies they employ in learning and their confidence and self-esteem, all of which impact profoundly on the quality of the learning achieved” (Broadfoot, 1996, p. 41). This suggests in turn that determining what constitutes ‘good’ assessment is a particularly crucial point. The knowledge that assessments come not only in a range of forms, but with different purposes and underlying philosophies, renders this matter akin to determining ‘fitness for purpose’ (Gipps, 1994).

Assessment is moreover a context-bound activity. “That which is taken for granted in one context can be puzzling, even unthinkable, in another” (Black, 1998, p. 21). It follows that the particular assessment techniques chosen at any particular time are determined by the respective national education systems that are each uniquely situated in a historical and contemporary social context (Broadfoot, 1996). As assessment practices reflect and reinforce the often conflicting values embodied in education systems, they thus need to be seen in relation to the wider societal and inter-societal forces acting upon them in order to be understood (Broadfoot, 1996). It must also be realised that as assessment practices are so deeply embedded in particular historical and social contexts, they cannot be easily changed (Black, 1998).

1.3 The Research Questions

1.3.1 Presenting the Research Questions

If someone had to ask me to describe in a nutshell what my research is all about, I would probably say that *it explores the Junior College PMI teachers’ classroom assessment practices*. This statement makes it clear that whilst the initial link with the

PMI option remains, the focus is on teachers and their assessment practices rather than the option itself. More importantly, this statement practically defines and delimits the boundaries of my study. For it signals unequivocally that the study concerns a particular aspect (i.e., classroom assessment practices) of a particular group of people (i.e., PMI teachers) within a particular context (i.e., they teach at the Junior College in the years that followed the introduction of the Intermediate Level courses). Given this specificity, my research problem has the ingredients to constitute a ‘case’ (see section 4.2.1) that is reflected in turn in the four main research questions that I eventually formulated and set out to answer when the study became more focused. These are:

1. How do PMI teachers view assessment?
2. What assessment practices do PMI teachers use inside their classrooms?
3. How do PMI teachers explain their classroom assessment practices?
4. What are the implications for learning of the PMI teachers’ views on assessment and their classroom assessment practices?

These questions are built on the understanding that the ‘phenomenon’ under review and the ‘case’ are not one and the same thing (see Hammersley & Atkinson, 1995). Indeed, whilst my phenomenon is Teachers’ Classroom Assessment Practices (TCAP) – which are taken to include all the initiatives or actions undertaken or encouraged by the teacher in order to gain understanding into, to improve upon, to evaluate and to communicate about the teaching-learning situation within his or her classroom – my case involves viewing this phenomenon in relation to students’ learning.

1.3.2 Broadening out the Underlying Issues

I have sought to base my understanding of TCAP and their implications for learning on a variety of data sources. The resulting information can be classified under one of the following three categories:

- Views: Data in this category concerns – (i) what teachers understand by classroom assessment; (ii) the importance that teachers attach to assessment inside their classrooms; (iii) how teachers see their role, and that of students and others in classroom assessment; (iv) how teachers regard assessment training for teachers; (v) how teachers view classroom assessment in relation to teaching and learning;

(vi) whom teachers see as the beneficiaries of classroom assessment and in what manner; and (vi) how teachers perceive the quality of their classroom assessment practices.

- Practices: Data in this category concerns – (i) the assessment training received by teachers; (ii) the sources and forms of assessment used by teachers; (iii) the involvement of colleagues and students in classroom assessment; (iv) the assessment tasks used by teachers; (v) the teachers’ recording of, response to and communication of classroom assessment results.
- Explanations: Data in this category concerns – (i) how teachers account for the origins of their classroom assessment practices; (ii) the classroom assessment constraints identified by teachers and their effects on assessment practices; and (iii) how teachers would like to assess in an ideal, constraint free environment.

In addition, my awareness that assessment practices are context situated has encouraged me to gather additional data from teachers that serves to map out further the encompassing background of this study. This data – which concerns a variety of teachers’ professional experiences, perspectives and practices – relates to professional training, teaching experiences, the national and school assessment systems, the school’s mathematics department, the PMI option (including its students), and PMI lessons. This background knowledge, apart from facilitating my understanding of TCAP and their implications for learning, also helps in the process of making sensible suggestions as to how classroom assessment can actually improve learning. The grounding of TCAP within the teachers’ personal, school and national realities is particularly essential to this study, for it is only then that one can truly explore the impediments and prospects of having classroom assessment at the service of learning.

1.4 The Structure of the Thesis

The thesis opens with an invitation to the reader to join me on a research journey. This invitation, which serves to establish the parameters of the study, practically sets the ball rolling into the thesis proper that contains the five parts described below:

- Part One – Setting the Scene: Part One consists of Chapter 1 and Chapter 2. The idea behind Chapter 1 is to acquaint the reader with ‘Michael the researcher as a person’ before proceeding to present the research area and questions. By giving indications of who I am and what I stand for, I hope that the reader can better understand why I chose this particular research, the manner in which I have conducted it, and what I would like to achieve through it. Chapter 2, which provides the background to the research, includes information on the Maltese Islands and their educational system, the local 18+ Matriculation Certificate Examination of which the PMI option forms part, and the Junior College. This chapter also lays out a policy rationale for change in Maltese education in relation to the emerging focus within it on formative assessment and curriculum breadth.
- Part Two – The Literature Review: Part Two includes an introduction in addition to Chapter 3. Apart from giving a critical account of the literature, these serve to position myself in the field. In the introduction, which delineates the links between assessment and learning, I highlight the growing awareness that assessment has the potential to raise educational standards by promoting learning. I then explore in Chapter 3 educational assessment inside the classroom with particular reference to the paradigmatic shift from an examination culture to an assessment culture and the implications that this has had on the role of the teacher as classroom assessor. Prominence is also given to empirical studies that look at the teachers’ world of assessment and to accompanying reflections on how this reality compares to the canons of the emerging, alternative assessment paradigm.
- Part Three – The Methodology: Part Three spreads over Chapters 4 and 5. In Chapter 4, which explores the variety of issues that have guided the study, I present my philosophical, methodological and ethical positions, and elaborate on the important practical considerations that followed from these value positions – namely, the research methods employed. On the other hand, Chapter 5 focuses on the implementation of the study. The areas discussed here include the selection of the research site, gaining access to the site, approaching the participants and gaining their informed consent, and the data collection, analysis and interpretation phases. A cautionary note to the reader as to how the research findings are to be interpreted brings Chapter 5 to a close.

- Part Four – The Data Analysis: Part Four consists of Chapters 6 and 7. Throughout these chapters, apart from presenting the analysis of the data collected from the participants, I also refer to myself in my capacity as a practising PMI teacher. These self-references are mostly given inside bordered areas under the heading of ‘Michael’s Story’. Whilst Chapter 6 serves to present the participants and to highlight the characteristics of the ‘typical PMI lesson’, the main purpose of Chapter 7 is to present the classroom assessment practices of the participating teachers. This portrayal of TCAP is however embedded within the teachers’ wider experiences of and perspectives on matters related to classroom assessment. The aim is to give the reader not only an understanding of how teachers assess inside their classrooms and why, but also of where they stand on a number of related issues.
- Part Five – Discussion and Conclusions: Part Five consists of Chapters 8 and 9. Here I discuss the research findings presented in Part Four and explore ways in which classroom assessment can be better made at the service of learning in the light of the present understandings. In particular, Chapter 8 provides a focused summary of the resulting TCAP, examines the implications for learning of these practices, and offers insights into the resulting quality of classroom assessment. Then, in Chapter 9, teachers’ explanations and justifications for their classroom assessment practices are analysed as part of a strategy to identify how classroom assessment can be better aligned to assessment for learning in spite of the many existing constraints and difficulties.

The thesis comes to an end with a brief account entitled ‘My Journey as a Teacher Continues ...’. As I recount here my life at the Junior College once the fieldwork phase was over, I lay particular emphasis on my efforts to put into practice, especially at school and national levels, what I have learned so far in my quest for understanding and knowledge. For whilst as ‘Michael the researcher’ I have chosen to privilege knowledge over action, as ‘Michael the teacher’ I continue to be greatly interested in the educative potential of my research (see Hammersley, 2003).

CHAPTER 2

The Research Background

2.0 An Introduction to the Maltese Islands

The Maltese archipelago (or Malta, as it is more commonly known) is strategically situated in the centre of the Mediterranean Sea. Malta consists of two main islands – Malta (which is by far the largest and most populated island in the archipelago) and Gozo (a distant second). With a total surface area of 316 km² and a population of slightly more than one-third of a million, Malta is one of the most densely populated countries in the world. The central location of the islands, combined with the excellent harbour facilities, has strongly influenced Malta's history. Malta has in fact been virtually occupied by all the dominant nations in the Mediterranean region, the last being the British Empire. Malta gained independence in 1964, became a republic in 1974, and joined the European Union (EU) in 2004.

Although the national language is Maltese, Maltese and English are both recognised as official languages and are used interchangeably throughout the whole educational system up to university level. Even though educational success depends on proficiency in both languages, it is the student's ability to absorb a system of instruction through English that really makes the difference (Zammit Mangion, 1988). For not only are textbooks and examinations predominantly in English, but teaching is also carried out in English in a number of subjects, especially when foreign or non-Maltese speaking students are in class. This high dependence on English condemns local students who are not well versed with this language to scholastic failure (Borg, Camilleri, Mayo & Xerri, 1995).

2.1 The Framework and Structure of the Maltese Educational System

2.1.1 The Framework

The Education Division is the government's administrative arm entrusted with the execution of the government's educational policies. Schools in Malta are regulated by

a National Minimum Curriculum (NMC) of studies – the 3-16 NMC (see Ministry of Education, 1999) for institutions offering pre-primary, primary and secondary education, and the Post-Secondary Level NMC (1991) for post-secondary institutions. Schooling is generally co-educational at the pre-primary and primary levels, single-sex at the secondary level, and again co-educational from the post-secondary level onwards. Schools are either state or private institutions (which are run either by the church or an independent organisation). All students attending post-secondary and undergraduate university courses are awarded state maintenance grants on a progressive scale that are subject to regular attendance and making satisfactory progress. Until the late 1970s, pre-service teacher education in Malta consisted either of a two year (later extended to three) full-time course leading to a Teacher's Certificate or a one year full-time Post-Graduate Certificate in Education (PGCE). In 1978, pre-service teacher training passed in the hands of the newly founded Faculty of Education at the University of Malta. The Faculty currently runs two pre-service programmes – the four-year full-time BEd degree and the one-year full-time PGCE course.

2.1.2 The Structure

School attendance is compulsory between the ages of 5 and 16 – a total of 11 years of schooling, during which mathematics is a compulsory subject, that guarantee universal primary and secondary education. At the end of the primary cycle, students may sit for the 11+ state qualifying examinations (i.e., all successful students are guaranteed a placing in one of the state academically oriented secondary schools) and the 11+ highly selective examinations (i.e., only a limited number of places are offered) organised jointly by a number of church schools. Mathematics is one of the subjects included in both examinations. Unless they attend one of the private schools that offer both primary and secondary education, students who fail (or do not sit for) these 11+ entrance examinations are placed in one of the less academically oriented state secondary schools.

At the end of the secondary cycle, students may sit for the 16+ Secondary Education Certificate (SEC) examinations organised by the local Matriculation and Secondary Education Certificate Examinations Board (MATSEC). Mathematics is the second

most popular subject taken at SEC level. Depending on their aspirations and qualifications (generally SEC or equivalent), students may then join one of the available post-secondary institutions. At present, approximately 55% of secondary school leavers enrol in a full-time post-16 study programme (Ministry of Education, 2001). Post-secondary institutions offer either 'academic' or 'vocational' programmes. The academic route is linked to pre-university (or sixth form) and 'secondary completion' courses. Apart from the Junior College, sixth form education is available in four other institutions (three of which are private).

In line with the international access revolution (see Broadfoot, 1996), education statistics reported by the Central Office of Statistics (1997) and the National Statistics Office (2001) show that the number of Maltese students choosing to continue studying beyond the compulsory years has increased significantly over the past years. For instance, about 71% more students followed full-time post-secondary courses during the 1999-2000 scholastic year than during the 1988-1989 scholastic year. But this rapid growth in student numbers at post-16 level has been mainly in favour of the academic courses. The student population growth at tertiary level has been even more remarkable. In fact, the number of students following day courses at the University of Malta (the only university on the island) more than tripled between the 1988-1989 and the 1999-2000 academic years, and there is a growing trend of students joining part-time evening courses.

2.2 Certification and Assessment in Malta

2.2.1 Certification

The MATSEC Board organises certification examinations at the end of the secondary and post-secondary 'academic' cycles. These examinations have practically replaced the traditional Ordinary Level (O-Level) and Advanced Level (A-Level) examinations offered by the English examination bodies to foreign candidates (Ventura & Murphy, 1998). At both levels, certification is mostly based on summative examinations, as coursework is only included in a few selected subjects and is kept as low as 15%. None of the MATSEC mathematics examinations has an element of school or classroom-based assessment.

2.2.1.1 At SEC Level

All 16+ SEC examinations follow the principle of differentiation that stemmed mainly from a concern with positive achievement and producing examination papers that would allow all students “to show what they know, understand and can do” (Gipps & Stobart, 1993, p. 83). SEC examinations consist of two two-hour papers (see Secondary Education Certificate Examination Regulations, 1995). Whilst Paper 1 (which only covers the ‘core’ syllabus content) is common to all candidates, students have to choose on registration between Paper 2A (which is more demanding than Paper 1) and Paper 2B (which is less demanding than Paper 1). Paper 2A is designed for the more academically able students, and is targeted at those who expect high achievement and want to proceed to higher education in the subject (Sultana, 1998). Candidates sitting for Paper 1 and Paper 2A may qualify for grades 1 to 5, and candidates sitting for Paper 1 and Paper 2B may qualify for grades 4 to 7. Students wanting to proceed to a sixth form college need to obtain grades from 1 to 5 in a specific number of subjects (the list varies from one college to another).

2.2.1.2 At Matriculation Level

The Matriculation Certificate was launched in 1995 “to replace the English GCE A-Level system ... with an International Baccalaureate-type system, without compromising the opportunity for students to achieve A-Level standard in subjects required for further study” (Ventura & Murphy, 1998, p. 48). The matriculation system is thus meant to create the ‘right’ tension between curricular breadth and depth:

Malta’s educational policy is based on the premise that individuals are more likely to develop into mature persons if their studies cover both the Humanities and the Sciences areas. ... The Matriculation Certificate Examination will be offered in a wide range of subjects divided among a number of subject groups which are structured in such a way as to ensure that all candidates choose subjects from both the Humanities and the Science areas. (MATSEC, 1994b, p. 1)

The Matriculation Certificate Examination moves away from the notion of a single subject examination system (see Matriculation Certificate Examination Regulations, 1995). Students must instead sit for the examinations of the six linear courses leading to the Matriculation Certificate in one session (i.e., in May, with the possibility of

resits in September). Of these six matriculating subjects – which must necessarily span over both the humanities and the science areas – two subjects must be at A-Level and the other four subjects at Intermediate Level (I-Level) (see section 2.3). As from October 1997, Maltese students can only join undergraduate courses at the University of Malta if they hold the Matriculation Certificate.

2.2.2 Assessment

2.2.2.1 Assessment that is Traditionally Based on Examinations

The traditional British examination system, with its emphasis on summative assessment, was taken up lock, stock and barrel by the Maltese educational system and incorporated within its foundations and structure (Mifsud, 1991). Maltese society in fact uses examinations, most of which are of the written type, as the yardstick by which to measure an individual's educational, cultural and intellectual ability from the earliest years at school (Zammit Mangion, 1992). Examinations – most of which rely heavily on the testing of the knowledge of facts, and rarely stretch students in other categories such as comprehension, application, analysis, and evaluation (Sammut, 1994) – have consistently played a decisive role in local assessment. Calleja (1988) in fact speaks of an educational bureaucracy that is intent on judging students' success only by examinations, and Pollard (1998) comments from his experience inside local classrooms that Maltese students seem to be predominantly assessed by examinations and tests. It thus appears that assessment in Malta, far from being the vital pedagogic instrument in curriculum design and development under the control of the teacher, is used instead as an instrument to measure or encourage performance, and to allocate students in different streams and schools (Mifsud, 1991). This selection, stratification, channelling and exclusion of students by means of formal assessment occurs on an ongoing basis throughout the whole system, not just at the 11+, 16+, and 18+ examination stages described in sections 2.1.2 and 2.2.1.

2.2.2.2 Growing Uneasiness with Malta's Examination Culture

Over the past years, there have been various appeals to shift local assessment away from the purely examination-oriented tradition (e.g., Fenech, 1988; Malta Union of

Teachers [MUT], 1996; Mifsud, 1991; Post-Secondary Level NMC, 1991; Wain et al., 1995). These appeals were motivated by the negative effects that the local examination culture has on methods of teaching and learning, and the well being of all involved.

- Teaching and learning: With examinations practically driving the curriculum (Fenech, 1988; Sultana, 1997), teachers are teaching with examinations in mind, hurrying through the vast syllabi and applying methods of teaching that are more conducive to rote learning than learning – and this to the detriment of a holistic curriculum that should cater for the all-round development of students (Wain et al., 1995). The teacher notes, model answers and the working of examination papers are consequently at the fore of the classroom experience (see Sammut, 1994). Headteachers and teachers get so caught up in this examination frenzy that they cannot think about anything else, and the highly specific examination syllabi practically exclude them from giving scope to their imagination (Calleja, 1988). This inordinate emphasis on examinations leads to a ‘culture of competitive achievement’ that culminates in the ‘diploma disease’ (Wain et al., 1995).
- Examination-related stress: The local exam-oriented schooling process causes excessive stress amongst teachers, students and parents alike (Farrugia, 1994). Not only do students preparing for examinations have to give up the activities they enjoy most, but they also develop significant stress-related behavioural changes (Mansueto, 1997). And whole families live under continuous psychological tension as children’s failure in examinations is almost always judged as a failure of the family (Calleja, 1988). Self-esteem in Malta appears to be linked with examination performance. For we tend to value ourselves and others tend to value us on the basis of examination results – that is, value is attributed to achievement and valuelessness to examination failure (Chetcuti & Griffiths, 2002). The obsession amongst Maltese parents to have their children pass examinations is leading in turn many students to attend several private lessons after school hours (A. Camilleri, 1995).

The suggestions that were made from time to time as to how reformed assessment practices may tone down the emphasis on competitive achievement and become more beneficial for student learning (e.g., the introduction of continuous assessment and

student profiling – see Wain et al., 1995) culminated in the formulation of Principle 9 of the 3-16 NMC (Ministry of Education, 1999) (see section 2.6.2).

2.3 The Intermediate Level of the Matriculation Certificate

2.3.1 ‘Defining’ the Intermediate Level

The I-Level – which many have perceived as ‘an unwelcome animal of Maltese origin’ (Zarb Adami et al., 1999) – has probably been the single most ‘visible’ and ‘problematic’ feature of the new matriculation system. Some (e.g., Briscoe, 1998; Zammit Mangion, 1994) have linked this new level, loosely defined by MATSEC as ‘one-third of an A-Level’, to fears of ‘lowering standards’ at pre-university level. Such fears echo the situation in the United Kingdom where the persisting notion that A-Levels are the benchmark of academic excellence impedes the widening of the 16-19 curriculum (see Pound, 1998). When MATSEC (1994b) announced that “the Intermediate Level will roughly require a third of the study time estimated as necessary to reach Advanced Level after the attainment of the Secondary Education Certificate” (p. 2), it did not specify whether the ‘one-third’ referred to the amount or the level of the content, or both. This ‘vagueness’ eventually led to many official complaints directed at MATSEC from students, parents and teachers alike (Zarb Adami et al., 1999). On their part, the MATSEC I-Level syllabus and paper setters panels have had to come up with their very own interpretation of its meaning.

2.3.2 The Pure Mathematics at Intermediate Level (PMI) Option

Successive MATSEC PMI syllabus panels have interpreted the ‘one-third of an A-Level’ directive to apply to both the content spread and depth of the corresponding Pure Mathematics at A-Level option. The main characteristics of the PMI option are delineated in the introduction to its first syllabus:

[It is] a subset of the syllabus for Paper 1 of the Advanced Pure Mathematics Syllabus. It can be considered to be the ‘core within the core’ and represents those topics in Pure Mathematics which are essential for the understanding of any mathematics (whether pure or applied) beyond school-leaving level. (MATSEC, 1994a, p. 116)

MATSEC assesses the PMI option by means of a three-hour-long examination paper that consists of 10 to 11 compulsory questions. MATSEC statistics show that whilst the registration figures for this examination almost doubled between 1997 (284 candidates) and 2002 (522 candidates), the national pass rate (i.e., grades A-E) over the same period has fluctuated between a high of 87% in 1998 and a low of 60% in 2001. A pass in the PMI examination is an entry requirement to a number of undergraduate courses at the University of Malta.

2.4 The Junior College of the University of Malta

2.4.1 Rationale, Aim and Structures

The University of Malta had long been insisting that students were not being adequately prepared for tertiary studies (Wain, 1995; Zammit Mangion, 1995). The University sought to address this situation by setting up the Junior College in 1995 in conjunction with the introduction of the Matriculation Certificate. In reality, the founding of the Junior College meant that the Education Division literally handed over its main sixth form college to University. Even though the aim of the Junior College is “to initiate students who have completed secondary schooling in the methods of study appropriate to tertiary education” (section 2 of the first schedule of the G. F. Abela Junior College Regulations, 1995³), a number of students still attend this institution without any intention of joining the University, and a minority even attend primarily for the maintenance grant (Junior College, 2000a).

Within each of the Junior College’s 23 departments, the school administration appoints a subject coordinator following consultations with the staff. This post is for a period of four years, and is renewable. The University Council appoints the school principal (also for a four-year renewable period) from amongst the subject coordinators. Although the principal sees to the day-to-day running of the institution, the Junior College is actually governed by the College Board that meets at least once every three months. Appointed teachers at the Junior College, who typically have at least a bachelors degree, do not need to possess any pedagogical training (nor are they

³ Although the University sixth form college is popularly known as Junior College, its legal name is G. F. Abela Junior College.

given any once they join) or teaching experience. Their remuneration increases progressively in line with their level of academic qualifications. Each week, teachers have a maximum teaching load of 12 one-hour sessions and are available for four personal contact hours during which students may consult them on an individual basis.

2.4.2 Entry Requirements and its Impact on the Educational System

Unlike most other local sixth form colleges, the Junior College does not operate a selective entry policy to its two-year matriculation courses. It accepts all students who hold the minimum entry requirements. These are passes at SEC level (i.e., grade 5 or better, even if obtained from Paper 1 and Paper 2B – see section 2.2.1) in at least six subjects, English and mathematics included. In its short and chequered history (see Buhagiar, 2003), the Junior College has had a considerable impact on the Maltese post-16 academic sector. Education statistics show that whilst its student population is continually increasing (exceeding, according to the latest published statistics, 2400 by the 1999-2000 scholastic year), about three out of every four fully qualified Maltese students following sixth form studies attend the Junior College (Central Office of Statistics, 1997; National Statistics Office, 2001). MATSEC statistics covering the examination sessions from 1997 to 2002 indicate in turn that Junior College students account for about between 50% to 60% of the Matriculation Certificate Examination's registrations and passes. This 'numerical growth' has not however been reflected in its funding. According to the Junior College (2000a), inadequate government funding not only hinders its proper overall development, but also causes a deterioration of its academic and support services.

2.4.3 The Academic Calendar, Regulations and Teaching Methods

Although the scholastic year spreads over three terms, second year students in practice attend for only the first two terms as the Matriculation Certificate Examination is scheduled in the third term. Students attend the Junior College from Monday to Friday. With the exception of Wednesdays – when lectures break off early in the afternoon – there are nine lecture slots of one hour each per day starting at 8.00 am. The Junior College, which does not practise streaming or other forms of selectivity, prides itself of promoting methods of teaching and study considered to be consonant with

university preparation. The overriding idea is to have “the type of teaching and pedagogical practices that ... bridge the chasm between the modes of teaching and learning that characterise secondary schooling and those that characterise tertiary education” (P. Camilleri, 1995, p. 6). To achieve this, the Junior College employs mixed teaching modes – lectures with relatively large groups of students (up to 50) are balanced by small group teaching methods such as seminars, discussions, workshops, tutorials, practical sessions, and personal contact hours.

2.4.4 The Evaluation and Promotion System

The Junior College regulations lay down that “The evaluation, assessment and eventual progress of students is determined by their work and contribution, which need not necessarily be solely academic” (Junior College, 2000b, p. 12). In practice, this system combines elements of continuous (read ‘Assessment Marks’) and summative (read ‘end-of-first-year examinations’) assessment to record the students’ academic progress and to promote them from first year to second year. At the end of each term (i.e., on three occasions with first year students and on two occasions with second year students), teachers forward an Assessment Mark on each student to the school administration, which is later sent to the student’s home. These Assessment Marks are out of 100 and in multiples of 5. Teachers may apply whichever criteria they deem fit to arrive at their Assessment Marks.

The first year Assessment Marks form an integral part of the Junior College’s promotion system, as they make up for 30% of the total score for each subject. The remaining 70% of the total score is based on the end-of-first-year examination. In order to proceed to second year, a student must obtain at least 40% in each of the A-Level subjects, and an average of 40% of the global mark in the I-Level subjects. A maximum of 10 marks may be added to this I-Level global mark when a student participates in officially approved extra-curricular activities. Students are allowed to repeat each year of the two-year course just once. The Junior College keeps an ‘Academic Report’ on each student that consists of the Assessment Marks, the end-of-first-year examination results, and any official correspondence or report concerning the student. The school’s end-of-first-year examinations and teachers’ Assessment Marks have no weighting on the final outcome of the Matriculation Certificate.

2.5 The Department of Mathematics at the Junior College

2.5.1 The Physical Environment

The mathematics department consists of an L-shaped corridor that has, for most of its parts, rooms (i.e., classrooms or teachers' offices) on one side and windows overlooking one of the school car parks on the other side (see Figure 2.1). The mathematics teachers share the four offices (A, B, C, and D) in this corridor. Office groupings range from one person (i.e., the subject coordinator in office A) to up to seven persons in an office. All the classrooms have conspicuously bare walls and the furniture inside is minimal – a whiteboard, students' tables and chairs, and a table and a chair for the teacher. The standard classroom layout is rows of student tables facing the whiteboard at the front. Although classrooms vary in size, many of them are too small for most of the mathematics classes (see Coleiro, 2004).

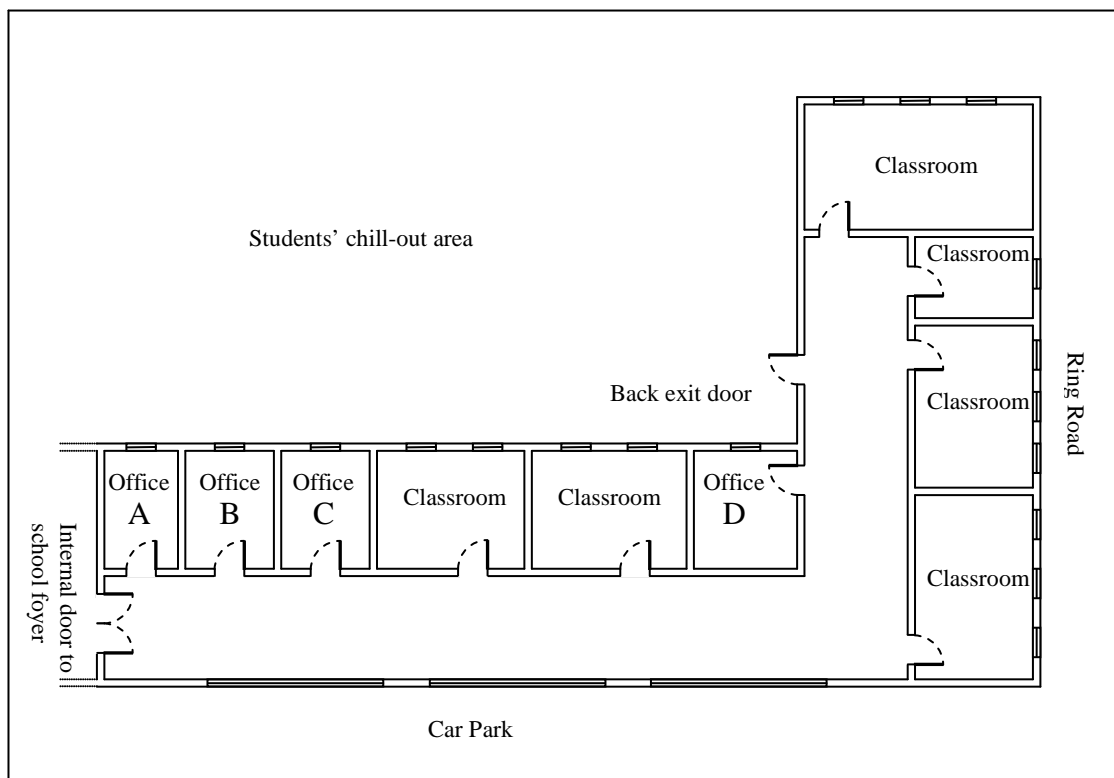


Figure 2.1: The layout of the mathematics department

2.5.2 Teachers and Operational Dynamics

Fifteen of the 20 full-time teachers at the mathematics department joined the Junior College when it opened in 1995, with the rest joining later. Ten of the initial 15 members had previously taught at the Education Division's sixth form college that the Junior College replaced. Table 2.1 indicates how the teachers are largely grouped in offices either by their pre-College affiliations or date of joining. Thus, whilst teachers in office C are the older and more experienced ones and teachers in office D are the youngest and least experienced ones, teachers in office B lie somewhere in between

Table 2.1: Teacher allocation inside offices

- **Office B:** All the teachers here joined in 1995, but most of them came from schools other than the replaced sixth form college.
- **Office C:** All the teachers here joined in 1995 and were also members of the mathematics department of the replaced sixth form college.
- **Office D:** All the teachers here joined at some point after 1995.

these two extremes. The teachers are predominantly male (only 6 females out of 20) and their ages span rather evenly from the late twenties to the late fifties. What follow are some of the more important operational dynamics within the mathematics department that have developed over the years:

- Staff Meetings: During the study, the department met at most on three separate occasions in any one scholastic year.
- Allocation of Classes: The norm within the department is for each teacher to have two classes that vary in level (i.e., one at A-Level and one at I-Level) and year (i.e., a first year class and a second year class). It is highly unlikely for a class to have the same teacher over the two-year programme.
- Syllabus Coverage: The MATSEC syllabus is divided into two parts – the first part is covered during the first year, and the second part during the second year. Teachers are free to organise the syllabus coverage as they please as long as the

respective parts of the syllabus are covered in time for the examinations (i.e., end-of-year for first year classes and MATSEC for second year classes).

- Classroom Assessment Practices: Teachers are free to organise their own classroom assessment practices, and there are no school or departmental criteria that guide the computation of the Assessment Marks.
- End-of-First-Year Examinations – Paper Setting and Marking: The subject coordinator sets by himself all the end-of-first-year examination papers, and only a few ‘trusted’ teachers are invited to review them in advance. Whilst the teachers are allowed to mark the scripts of their own students, no standardised marking schemes are made available and there is no departmental moderation of marking.

2.5.3 The PMI Option

A PMI class – which meets for three one-hour periods per week – retains the same students (at least those who get promoted) over the two-year programme. In Year 3 of the study, when the main field data was collected, the average class size for the PMI option was 39 students. Although the Junior College suggests to its prospective students that “a good knowledge of the SEC Mathematics syllabus is desired for students taking Intermediate Pure Mathematics” (Junior College, 1998, p. 28), students with grade 5 in SEC level mathematics (even if obtained from a combination of Paper 1 and Paper 2B – see section 2.2.1) are also allowed to take the option. Whilst the number of students taking the PMI option at the Junior College continues to increase (e.g., in Year 1 of study – 480 students out of a school population of 2222; Year 2 – 501 out of 2033; and Year 3 – 674 out of 2431), MATSEC statistics show that in the first six years of the Matriculation Certificate Examination, an average of 76% of the Junior College students who registered for the PMI examination obtained a pass.

2.6 Rationale for Change in the Maltese Educational System

2.6.1 The Current Situation

The Maltese educational system – which Ventura and Murphy (1998) define as ‘intrinsically inequitable’ – has traditionally promoted the ‘talented’ few at the expense of the rest (Darmanin, 1992). As a result, in spite of the recent increases in student numbers at post-secondary and tertiary levels (see section 2.1.2), Malta still lags behind most EU countries in this respect. Suffice to say that only 37% of Maltese 18-year-olds were still in education in 2000 compared to the EU-15 average of 75% (European Commission, 2005a), and in 2004 the number of Maltese aged 20 to 24 who had completed upper secondary education stood at 48% (the lowest in the EU) whereas the EU target for 2010 is 85% (European Commission, 2005b). In a system still characterised by the constant closure of opportunities from the very early years of primary schooling (Sultana, 1997), Maltese lower achieving students face humiliation and neglect, and normally end up fighting this ‘oppression’ by opting out or playing truant (Chircop, 1997). Not only are local students continually being ‘weeded out’ of mainstream education, but research also evidences that this is done for reasons not related to education (e.g., social class – see Sultana, 1995; date of birth – see Borg & Falzon, 1995, and Borg, Falzon & Sammut, 1995; and gender – see Darmanin, 1992). The net result is a highly differentiated educational system:

With regards to the world of Maltese schools, what is immediately striking is the practice of differentiation. There is differentiation between sectors (state, church, independent), and between schools within the same sector (streaming in primary schools, and between different types of secondary schools Differentiation is also practiced in the post-16 sector, with the Junior College, vocational institutes and establishments, and apprenticeship schemes having the status of separate tracks, with little connection between them. (Pollard, 1998, p. 174)

This ‘fractured’ reality – which has the support of the vast majority of parents and teachers (see Mercieca, 1997) – is regulated by means of a high-stakes examination system (see section 2.2.2.1). This leads to a form of schooling that

is solely geared towards the achievement of good academic results. Headteachers and teachers do not think of anything else; they become part of the bureaucracy intent on judging students’ success only by written examinations. (Calleja, 1988, p. 31)

In spite of the many calls for assessment reform in Malta (see section 2.2.2.2), examinations continue to dominate the assessment system. A recent survey on current assessment practices in local schools in fact concludes that these are still very much embedded within the traditional culture of examinations and testing, and that assessment is used mostly summatively (Grima & Chetcuti, 2003). Within this scenario, mathematics examinations feature prominently, as wherever streaming is practised, it depends to a large extent (especially at primary level) on students' performance on mathematics examinations. And students can only join one of the more academically oriented secondary schools (state or private) if they pass, amongst other subjects, the obligatory mathematics entrance examination. Again, success in SEC level mathematics is a prerequisite to join many post-secondary institutions, particularly the sixth form colleges. Even at university level, a mathematics post-secondary qualification is a special course requirement for many of the available undergraduate courses.

The local tendency to continually filter students as they move up the educational ladder is embedded in turn within school curricula that are often overloaded – a condition that, as Wain (1991) remarks, is likely to constrain teachers and schools to set different priorities than those officially recommended. *Tomorrow's Schools* (see Wain et al., 1995), a report that examines local educational practices and policies, in fact warns that the overloaded curricula, in conjunction with the local overemphasis on examination achievement, are promoting a culture of coverage as opposed to understanding and a classroom pedagogy aimed primarily at short-term targets. The report also notes a strong predilection for “traditional modes of teaching that privilege memory and the storage of information” (p. 45). The system moreover “tends to inhibit the child and train him to listen, obey, follow, learn and conform with set rules of thinking, feeling and behaving from the earliest stages, becoming even more oppressive and conformist (due largely to examination influence) as he/she mounts the scale” (Zammit Mangion, 1994, part two, p. 5). Instead of promoting independent and creative thinking, the system thus encourages a ‘culture of mimicry’ (Bartolo, 1997). Even the best students – who are very good at reproducing what they have read in books – find it difficult when faced with a real problem that requires them to analyse, to think, and to apply what they have been taught to a real situation (Bartolo, 1997).

2.6.2 The Vision Ahead

Providing students with high quality education is thought to be the key to enable Malta to continue to develop and prosper, both socially and economically (Chalmers et al., 2004). Successive Maltese governments have in fact been building on existing measures to create and support a wide range of learning opportunities for all age groups (Mallia et al., 2001). These initiatives, which basically seek to guarantee open access to learning, are based on the understanding that the Maltese are increasingly becoming aware that the process of learning is a lifelong and lifewide one, that learning makes them better persons, and that such a process entails continually gaining and upgrading one's skills (Mallia et al., 2001). Embedded within this vision lies the democratisation of the educational process – a shift from an emphasis on the education of all with a clear focus on the better students to quality education for all (see Principle 1 of the 3-16 NMC – Ministry of Education, 1999). At post-secondary level, these 'wider access' and 'quality education' policies parallel the two premises on which the Post-Secondary Level NMC (1991) is built. These are: (i) that formal education beyond the mandatory school age will become the norm rather than the exception in the very near future; and (ii) that albeit studies at the post-secondary level should signal the initiation of the 'process of real specialisation', this should not be achieved at the expense of severing studies from the 'main trunk of knowledge', as it is wise to have a common core of knowledge and skills for securing its very quality and relevance. But whilst the replacement in 1995 of the traditional three A-Level system by the Matriculation Certificate Examination sought to address the desire for 'a broad general education' at pre-university level without compromising future specialisation (see section 2.2.1.2), formal education beyond compulsory schooling is still a long way from becoming the norm (see section 2.6.1).

This evolving scenario is complemented by ongoing efforts to establish learning as an active and relevant process (see Principles 3 and 4 of the 3-16 NMC – Ministry of Education, 1999). This recognises that students are capable of transforming and personalising new knowledge and that they appreciate learning more when it has application and relevance for everyday life. The embedded constructivist understandings that students need to have access to the goals of their learning and that teachers need to know what sense students are making of their learning experiences

have rendered it necessary in turn for Malta to reform its traditional exam-oriented assessment practices (see section 2.2.2). The current initiatives towards ‘a more formative assessment’ in local schools (see Principle 9 of the 3-16 NMC – Ministry of Education, 1999) are meant in fact to address this pressing need for the sake of both students and their teachers. In particular, the 3-16 NMC envisages that whilst the process of assessment in the first three years of the primary school will be entirely formative, summative and formative modes of assessment will be linked together as from then onwards. The prudent approach to assessment reform expressed in the 3-16 NMC and subsequently being adhered to in the implementation process echoes a general consensus crystallised by the MUT (1998) position that “Any alternative ways of assessment must be introduced gradually and only after careful planning” (p. vi). This sense of caution probably recognises the fact that removing examinations from the centre of assessment would entail what Mifsud (1991) describes as a ‘cultural shift in the Maltese society’. For this would necessitate the creation of a cultural climate in which the community would be made aware of the true aims of education and the place that examinations have in the whole educational enterprise (Fenech, 1988).

2.6.3 Positioning the Thesis

The people who help define educational policy in Malta – be they politicians, Education Division officials, teacher educators or the Malta Union of Teachers – agree in principle on the key issues of quality education for all, the need for more students to continue with their education beyond mandatory school age, curricular breadth at pre-university level, and the need to render assessment practices more formative at school level. This study thus explores the classroom assessment practices of a group of sixth form mathematics teachers operating within a context that is supposedly guided by this vision. So far, though, the only tangible changes for these teachers are the ever-increasing number of students continuing with their pre-university studies (albeit still short of the set target), the introduction of differentiated papers as SEC level, and curricular breadth at sixth form level with the launch of the Matriculation Certificate. An opportunity consequently arises to understand TCAP in relation to how teachers experience the everyday realities of their school and classroom contexts that are embedded in turn within the demands and expectations formulated from outside school in official policy documents, legislation and other forms of communication.

Part TWO

THE LITERATURE REVIEW

INTRODUCTION

Learning and Assessment

This study is primarily concerned with exploring Teachers' Classroom Assessment Practices (TCAP) in relation to students' learning (see section 1.3.1). It is a focus that reflects the increasing awareness amongst leading academics about the crucial link between assessment (as carried out in the classroom) and learning and teaching (see Assessment Reform Group [ARG], 1999). This builds on the recognition that learning goals, teaching activities, learning processes and assessment procedures coexist in a system of interrelationships in which all four components are in dynamic tension or balance (see Cumming & Maxwell, 1999). Because assessment both reflects and communicates deeply held convictions about how students learn, what students should learn and why (Denvir, 1989), changes and innovations in theories of learning (which influence in turn conceptions of appropriate instruction) and assessment cannot but happen concurrently (Cizek, 1997). Particularly important to this study – and what this introduction to the Literature Review sets out to achieve – is the examination of what led to our shift in understanding of the learning process from the behaviourist to the constructivist model, and the implications that this 'revolution' has had for assessment. This would then serve to better identify the classroom assessment practices that are in line with and facilitate the current reform efforts in Malta (see section 2.6).

Psychological theories of learning, which date back to the early part of the 20th century, are underlined by two linked assumptions – decomposability and decontextualisation (Gipps, 1994). Decomposability is based on the notion that learning of complex competencies – just like 'building blocks' – can be broken down into discrete skills learnt separately by developing individual stimulus-response bonds. In what clearly follows the canons of behaviourist theory, learning is seen to be linear and sequential, with complex understandings only occurring by the accumulation of elemental, prerequisite learnings (Shepard, 1991). This learning model, which is characterised by practice, consequently precludes moving to higher levels until the prior level has been mastered, and rests on the idea that repetition is the only way to remedy deficient skills acquisition (Shepard, 1991). I would however argue like Gipps (1994) that theories built on the assumption of decomposability, apart from never

articulating clearly how complex skills can be developed later, work against the development of problem solving or thinking skills as learning within this framework tends to focus on separate skills. Behaviourist learning models moreover fail to recognise the current understanding that practising higher order skills can actually help to develop or strengthen 'basic skills' (see Gipps, 1994). Decontextualisation, on the other hand, is based on the premise that "each component of a complex skill is fixed, and that it will take the same form no matter where it is used" (Resnick & Resnick, 1992, p. 43; cited in Gipps, 1994). This builds on the traditional notion of knowledge as an "integral, self-sufficient substance, theoretically independent of the situations in which it is learned and used" (Brown et al., 1989, p. 32). But situated cognition theorists now challenge this separating of what is learned from how it is learned and used, and argue instead that the activity in which knowledge is developed and deployed is an integral part of what is learned (Brown et al., 1989). This understanding that situations co-produce knowledge through activity (which also contributed to my distancing from the ontological position of 'objective knowledge' and thus paved the way for the methodology used in this study – see section 4.1) lead me to argue that skills, knowledge and facts cannot be learned in isolation and then used in any context.

Black (1998) contends that an atomised approach to learning – which decomposability and decontextualisation imply – emphasises learning by rote, of small pieces of information without the understanding that interrelates them, and of fixed rules and procedures. These rules and procedures will consequently be grasped only as tactics, without the strategic overview needed to give them significance and to guide their application. This form of learning is not helpful to students who will be expected to be flexible, adaptive and able to change in response to the rapidly developing and complex technological society (Denvir, 1989). Although there are some things that are probably most efficiently learned by rote (e.g., number bonds, spellings, and multiplication tables), the exponential increase in the amount of factual information in recent years and for the foreseeable future, coupled with the rapid changes in the nature of employment, indicate that there should be far greater emphasis on learning that can be transformed and applied to new circumstances than on learning facts and procedures applicable only in situations closely similar to those in which they were learned (Harlen & James, 1997). A simple but powerful way of characterising these two approaches to learning – the former leading to understanding and the latter to rote

memorisation – is to distinguish between ‘deep learning’ and ‘surface learning’. Deep approaches involve an active search for meaning, underlying principles, structures that link different concepts or ideas together, and widely applicable techniques; surface approaches, in contrast, rely primarily on attempts to memorise course material, treating the material as if different facts and topics are unrelated (Marton & Säljö, 1976). A surface or shallow approach to learning – especially if used indiscriminately or habitually (see Harlen & James, 1997) – is clearly inappropriate because “Isolated facts, if learnt, quickly disappear from the memory because they have no meaning and do not fit into the learner’s conceptual map. Knowledge learnt in this way is of no use because it cannot be applied, generalized, or retrieved” (Gipps, 1994, p. 21).

Shepard (1991) identifies the ‘test-teach-test’ cycle as one of the fundamental principles that underlie a behaviourist approach to educational practice. This insistence on testing builds on the now discredited “assumption ... that one can specify and measure all important learning objectives, and furthermore that mastery on the test items implies mastery of the intended skills and concepts” (Gipps, 1994, p. 20). The accompanying belief that for any task there is a ‘true score’ that can be ascertained and reported for an individual test or examination⁴ taker (Torrance, 2000) accounts for the traditional assessment scenario that is primarily concerned with weighing up individuals and comparing them to others (see Gipps & Murphy, 1994). This dominant reality fails however to appreciate that as any assessment is based on a sample of the behaviour that interests us (see Nuttall, 1989), any educational measure can offer at best only a rough estimate of particular kinds of ability (Broadfoot, 1996). Lacking this awareness, the emphasis within the behaviourist tradition has been on producing quick, economic tests that have acceptable psychometric properties and are easy to mark, rather than assessment procedures that provide a useful picture of what students can do (Broadfoot, 1996). In particular, testing within the positivist tradition typically asks for demonstration of small, discrete skills practised in isolation that engages the learner in finding someone else’s correct answer rather than in personal interpretation and thinking (see Gipps, 1994; also Black, 1998). But whilst this may help us find out what students remember about what we think they should remember, we do not get to truth,

⁴ For the purpose of this study, the main distinction between a test and an examination is taken to be that whilst the former is planned and administered by the teacher within the confines of his or her classroom, the latter originates outside the classroom at either school or national level.

meaning, purpose or utility (Ellis, 2001). This leads me to argue like von Glasersfeld (1995) that the behaviourist school's success in eliminating the distinction between training for performance and teaching that aims at the generation of understanding has had unfortunate consequences for education:

[The behaviourist learning theory] has tended to focus attention on students' *performance* rather than on the *reasons* that prompt them to respond or act in a particular way. Reinforcement fosters the repetition of what gets reinforced, regardless of the acting subject's *understanding* of the problem that was posed, and of the inherent logic that distinguishes solutions from inadequate responses. Thus, training may modify behavioural response, but it leaves the responding subject's comprehension to fortunate accidents. (p. 4) (emphasis in original)

Within the dominant behaviourist teaching/learning/testing model, teachers focus on discrete skills and on decontextualised test items, offering over-practice in the hope of achieving mastery (Gipps, 1994). But albeit an 'explanation and practice' approach may seem to cover ground more quickly, progress is often illusory (and hence long-term learning elusive) as imitative methods usually develop dependency and a fragile fluency that is lost when practice ceases (Swan, 2001). Within this traditional approach – which contradicts the assumption that an important aim of education is to bring about learning with understanding (see Harlen & James, 1997) – teaching is accomplished by telling and learning by repetition, and the learner is viewed as a passive container waiting to be filled with knowledge, but possibly not receiving the knowledge because of a 'block' (see Denvir, 1989). The underlying assumption that knowledge can be transferred from teacher to student explains why traditional teachers are mainly concerned with getting knowledge into their students' heads (von Glasersfeld, 1989). Moreover, the embedded view of instruction as the delivery of information and the decoding of that information as the responsibility of students lies behind the 'mistaken' notion that teachers' responsibility basically ends when they have told students what they must remember to know and do (Sedlak, 1987; cited in Fang, 1996).

It follows that the purpose of assessment within this transmission framework is to determine the effectiveness with which the teacher communicates a body of knowledge to students (Brown, 1989). And the feedback it provides emphasises what has not been learned (Harlen & James, 1997) and just tells students – often too little and too late – how they have done on a test or in a course, not how they are doing as

learners (Cross, 1998). Nowadays, however, the desire to encourage deep learning strategies favours instead an assessment system that places “emphasis on understanding, transfer of learning to untaught problems or situations, and other thinking skills, evaluating the development of these skills through tasks that clearly must involve more than recognition or recall” (Crooks, 1988, p. 468). This shift in assessment parallels the new conceptions of learning that, as Glaser and Silver (1994; cited in Cizek, 1997) explain, became necessary following the failure of behaviourist theories to describe adequately complex processes of thought, reasoning and problem solving. In fact, recent work in cognitive and constructivist psychology comes as an alternative to the behaviourist, linear hierarchy model of learning (Gipps, 1994, 1996). It is now recognised that different students possess different ‘kinds of minds’, and that they consequently learn, remember and understand in different ways (Gardner, 1991; cited in Gipps, 1996). This overdue recognition of heterogeneity is built, as LaCelle-Peterson (2000) points out, on the notion that each learner presents a unique profile of abilities, accomplishments, characteristics and needs. Contemporary cognitive psychology, which has built on the very old idea that things are easier to learn if they make sense (Shepard, 1991), consequently supports the notion that

... understanding involves creating links in the mind and that ‘making sense’ of something depends on these links. Isolated pieces of information do not have links to existing mental frameworks and so are not easily retained in the mind. The identification and creation of links to existing frameworks depends on the active participation of the learner and on the familiarity of the context of the material to be learned. Understanding, in this view, is the process of construction and reconstruction of knowledge by the learner. What is known and understood will, of course, change with new experience and as new ideas and skills are presented to help make sense of it. (Harlen & James, 1997, p. 368)

In this new understanding of learning as a cognitive and constructive process, knowledge is not an external map that is transposed directly into the student’s head, but results from the organic process of reorganising and restructuring undertaken by the student as he or she learns (Gipps, 1994). We now believe that students learn best by actively making sense of new knowledge – making meaning from it and mapping it into their existing knowledge map/schema (Gipps 1994). Constructivist learning theories acknowledge that students are active in their learning – a notion that belies their traditional role as passive receivers of knowledge. The student is now seen as agent, the active constructor of meaning and knowledge who shares responsibility for

learning with the teacher (Murphy, 1996). But whereas learning in some versions of cognitive theory is almost completely a function of the learner's interpretations of events (Steadman & Svinicki, 1998), I favour the social constructivist perspective that knowledge is a product of dialogue and negotiation between teachers and students (Murphy, 1996). The understanding here is that "we learn from being part of and interacting within a social environment, and that individual construction of knowledge is derivative of its social construction" (Jaworski, 2002, p. 73).

It follows that, in a constructivist environment, the students are expected to engage in dialogue with each other and with teachers, and to validate their own understandings rather than merely accept transmitted views (Murphy, 1996). This implies that students should be willing to consider each other's solutions and to be prepared to accept better solutions without an *a priori* acceptance of the teacher's view (Seegers & Gravemeijer, 1997). Improvement in learning is seen to depend on students coming to hold a concept of quality roughly similar to that held by the teacher, and on their ability to draw on a range of strategies to close the gap between their actual performance and the standard they are aiming for (Sadler, 1989). The emphasis is now on students as conscious decision-makers whose learning is based on personal commitment and deep consideration. The concept of agency by students – which rests on the premise that it is the student who constructs meaning out of the opportunities that the school offers – links educational progress to the need for students to gain an explicit understanding of what they know and how they come to know it (Murphy, 1996). Towards this end, students need to develop what von Glasersfeld (1989) calls 'operative knowledge' that, contrary to the traditional knowledge used for associative retrieval of a particular answer, is knowledge of what to do to produce an answer.

These roads to deep learning involve thinking about the meaning of what is being learned – a metacognitive process (which is basically the second-order practice of 'thinking about thinking') that helps the learner to plan, monitor, orchestrate and control his or her own learning through a variety of self-awareness processes (Gipps, 1994). But as metacognitive understandings do not just happen, the teacher has to help his or her students acquire the necessary skills and experiences. When teachers progressively turn over metacognitive functions to their students, students start to appreciate what it means to learn and gain awareness of their own learning strategies

and efficiency – this is when learning can turn into an intentional process rather than incidental (Gipps, 1994). But if the teaching and development of higher-order skills (e.g., application of knowledge, investigation, analysing, reasoning and interpretation) are to be encouraged, assessment needs to reflect such qualities (Gipps, 1994). Instead of using an assessment model that atomises knowledge and informs about recognition or recall of facts, we need to assess the level and complexity of understanding (Gipps, 1994). In particular, assessment facilitates metacognition when students are active in their own assessment, and when they see assessment as a moment of learning in the light of an understanding of what it means to get better (see Black & Wiliam, 1998).

Once knowledge and competence are recognised as products of the individual's conceptual organisation of his or her experience, the teacher's role will no longer be to dispense 'truth' but rather to help and guide the students in the conceptual organisation of their experiences (von Glasersfeld, 1989). This involves providing students with authentic activities that are meaningful and purposeful from their perspective, and that allow them to apply and develop their understandings in explicit relation to others (Murphy, 1996). The teacher also needs to orchestrate discussions around issues that are significant in view of the envisioned learning trajectories (Seegers & Gravemeijer, 1997). A discussion-based teaching approach built on the sharing and renegotiation of ideas in an atmosphere of mutual trust may appear slow initially, but learning becomes meaningful, connected and stable over time (Swan, 2001). In this scenario, the teacher becomes a facilitator who is alive to the shifts and turns in students' thinking and who encourages students to build on their relevances (Woods, 1990). This new interactive relationship between teacher and learner leading to learning was given prominence by Vygotsky in his 1962 seminal publication *Thought and Language* (Black, 1999). His theory of cognitive development rests on the key concept of 'internalisation':

Vygotsky argues that all higher psychological processes are originally social processes, shared between people, particularly between children and adults. The child first experiences active problem-solving activities in the presence of others but gradually comes to perform these functions independently. The process of internalization is gradual; first the adult or knowledgeable peer controls and guides the child's activity, but gradually the adult and the child come to share the problem-solving functions, with the child taking initiative and the adult correcting and guiding when she falters. Finally, the adult cedes control to the child and functions primarily as a supportive or sympathetic audience. (Brown & Ferrara, 1985, pp. 281-282)

It is within the context of this gradual internalisation of cognitive activities that Vygotsky (1978) introduced his concept of the Zone of Proximal Development (ZPD). The ZPD refers to the gap between what the learner can do on his or her own and what he or she can do with the help of others. The process of support and guidance offered by the teacher to help the student to perform at a higher level is known as ‘scaffolding’ – the teacher offers support that is gradually removed as the student becomes competent at that level (Gipps, 1994). The scaffolding metaphor signals that even though the teacher provides the scaffold for the building, the building itself can only be constructed by the learner (Black, 1999). In this supportive role, the teacher has to discern the potential of the student to advance in learning, so that the activities presented, instead of being either too trivial or too demanding, fall within Vygotsky’s ZPD area of appropriate and productive challenges (Black, 1999). The use of scaffolding understandably requires the teacher to be aware of individual students’ personal needs (Murphy, 1996). One function of assessment would then be to help identify this zone accurately and to explore progress within it. This implies moving assessment beyond the traditional static model of what is known towards a more interactive model that looks at learning potential in Vygotskyian terms (Gipps, 1994). This is not potential in a ‘static’ sense (i.e., it is determined and cannot be changed), but potential that is elastic and highly responsive to adult support and teaching. As such, assessment not only indicates what the student knows and can do, but also what he or she nearly can do (Gipps, 1994).

Current work in cognitive science for the assessment of student learning (as opposed to performance) suggests that we need to focus on the models that students construct for themselves and their understandings. “The challenge, then, is to find out enough about student understanding to design performances that will reflect these different understandings and then to design assessment techniques that can accurately reflect these different understandings” (Wilson, 1992, p. 125; cited in Gipps, 1994). This calls for skilfully directed assessment that reveals important aspects of learning and lays down the foundations for further growth and accomplishment. The recognition that assessment, as part of education, must be about promoting learning and opportunities (Brown, 1990) in turn underpins the current reform efforts aimed at shifting the dominant assessment paradigm “from a testing and examination culture to an assessment culture” (Gipps, 1994, p. 1).

CHAPTER 3

Teachers and Classroom Assessment

3.0 The Need to Move Away from the Traditional Assessment Paradigm

In the introduction to the Literature Review I argued that the notion of ‘assessment’ is starting to move away from the traditional vision of tests, examinations, selection, grades and marks in response to our growing understanding of learning as a meaning making process in which much depends on the learner’s constructions of his or her own experiences. But this emerging emphasis on assessment that is explicitly designed to promote learning (see section 3.1) is, according to my reading of the situation, also a reaction by ‘enlightened’ people in the educational sector to the multiple ways in which the still ubiquitous presence of the traditional forms of assessment – namely, tests and examinations – in contemporary educational systems (see Broadfoot & Black, 2004) affects negatively the teaching-learning environment. With the general public, however, the examination system continues to enjoy pervasive support (Broadfoot, 1996). This continued social acceptance appears to be linked to the persisting notion that examinations, which are believed to provide fair, objective, reliable and precise measures of achievement that everyone can easily understand in a manner that is not administratively taxing, promote equal opportunities (see Broadfoot, 1995; Brown, 1990; Gipps & Murphy, 1994). Although these ‘qualities’ have been by now either discredited (see Broadfoot, 1996; Gipps & Murphy, 1994) or adequately repositioned (see Murphy & Torrance, 1988), the sustained popular belief in them is making it hard to change existing educational systems in which the ultimate scope has traditionally been to select the most ‘talented’ for the most educational investment in the mistaken conviction that these would later be able to ‘put most back’ into the economy (see Torrance, 1995). Whilst efforts, at least in developed countries, to encourage students both to stay on longer within formal education and to equip and motivate them to keep coming back to education throughout their lives (Broadfoot & Black, 2004) attest to the economic foolhardiness of the ‘human wastage’ that underpin of such systems (Malta is becoming increasingly conscious of this – see section 2.6.2), there are, in my view, even stronger educational arguments, which go beyond the direct link between learning and assessment, as to why traditional assessment practices need to change.

For a start, the traditional privileging of assessment for selection has led to curricular choices that result in inappropriate course for the majority in order to satisfy a minority who wish to pursue university courses (Denvir, 1989). Other than this, the traditional system of testing and examining has severe limitations that, according to Desforges (1989), can be grouped under three categories:

1. A narrow assessment: Traditional examinations are limited to testing those skills and bodies of knowledge that can be appraised in a short, fixed period for all candidates at a particular point in time. This means that only a very narrow range of qualities (i.e., the academic skills, using paper-and-pencil techniques) is assessed. There is thus a frequent mismatch between the curriculum and examination content.
2. A sudden-death assessment: To allow comparability and hence credibility, all candidates in a given subject are examined at the same time. This means that they get credit only for what they do in that brief examination period – an assessment process that effectively ignores their work during the study course. Although those who work hard in the course are better set to do well in an examination, any error of judgement by a candidate in the examination would have catastrophic results.
3. A non-informative assessment: The final grade (or score) is related to the achievement of other candidates – for example, a grade says that, on the day, a candidate with grade C scored more points than those with a D or less, and fewer than those who graded B or better. This means that the result only informs how well a candidate did in the broad league table of all the others who sat for the examination. The grade does not define what a candidate knows and can do – there are no descriptions of specific and changing levels of attainment.

Particularly relevant to the present study is the limiting effect that testing has on teaching. The situation is such that the more significant an assessment is, the more likely that teachers will concentrate on teaching what that assessment measures to the detriment of the other untested skills and activities (Gipps & Stobart, 1993). Given the political and social uses to which the test scores of external assessment are subjected, it is not hard to understand why this form of assessment is generally recognised as a tool that standardises and controls what teachers do (Smith, 1991). It follows that

... although testing is usually considered to be a means of measuring qualities that are already present in a person, in actuality tests often *produce* the characteristics they purport to measure. The individual in contemporary society is not so much described by tests as constructed by them. (Hanson, 2000, p. 68) (emphasis in original)

I find this ‘construction’ of particular concern in view of the ‘narrowness’ that characterises traditional testing (see Desforges, 1989). Also of concern is the realisation that the knowledge that test scores are the sole means of describing and judging schools produces in teachers the determination to do what is necessary – which translates itself in a ‘testlike’ teaching programme – to avoid being publicly embarrassed and humiliated (Smith, 1991). Many teachers have in fact been known to even engage in questionable, educationally indefensible practices in order to increase their students’ test scores (Popham, 1991).

The first victims of a teaching programme that tends to focus on what is assessed – better known as ‘teaching to the test’ or ‘curriculum backwash’ – are the teachers. By holding back from using methods that do not conform to traditional testing formats (e.g., exploration and discovery), teachers deskill themselves by losing their capacity to use them or to imagine them as possibilities (Smith, 1991). The other direct victims are the students because, as Schoenfeld (1999) argues, this type of teaching leads to a set of skills that have little to do with deep competence. The fact that examinations have always emphasised recall of factual knowledge with a heavy reliance on memory and rote learning (Gipps & Stobart, 1993) leads teachers to coach their students in a narrow range of test-taking skills rather than encourage them to teach a broader range of higher order competencies and understandings (Torrance, 1995). As a result, students develop problem-solving strategies that help them pass examinations without reaching the intended learning goals (Brookhart, 1999) – which shows how misleading educational success as generally understood can be. Although improvements in examination results without any real improvement in educational quality have no real value (Torrance, 1995), for most students passing examinations is the only purpose of being in school (Broadfoot, 1996). The system is such that examination results command greater esteem than those of other forms of assessment (Black, 1998). Students, like most of society, place such great value on what is assessed (Bryant & Driscoll, 1998) that should the teachers decide otherwise, they are known to settle their dilemma between the conflicting messages coming from instruction and assessment,

especially if it is for high stakes, by assuming that the goals of assessment are the ones that count (see National Council of Teachers of Mathematics [NCTM], 1995).

Labelling – which occurs when test scores and examination passes determine ways of thinking about students – is another danger of traditional assessment (Gipps & Stobart, 1993). This phenomenon gives rise, in my view, to two important issues. First, the aura of objectivity enshrined in the traditional assessment model, which is built on winners and losers, projects the notion that competition within it is open and fair, and that consequently losers have only themselves to blame (Broadfoot, 1996). This positioning of blame is however misguided given that a substantial body of research testifies that tests (on which the ‘race’ is mostly based) do not even necessarily do what they aspire to do, that is, to provide objective, reliable evidence of the attainments measured (Broadfoot, 1996; see also Black, 1998; NCTM, 1995; Shepard, 2001 [cited in Schoenfeld, 2002]; Tolley, 1989). In particular, when the assessed domain is very wide, “any test set in a finite time could only test a small fraction of the content and skills that characterize expertise ... With such a small fraction, it would not be legitimate to claim that the result is a measure of attainment” (Black, 1998, p. 65).

The second issue with labelling is that it can affect teachers’ views about what students are capable of doing (Gipps & Stobart, 1993). In particular, it can set an unconscious limit on what students are perceived as being able to do. What I find most disturbing about this is that teachers are known to categorise students into ‘bright’ and ‘dull’, ‘able’ and ‘less able’ without any objective evidence of performance (see Claxton, 1994; cited in Broadfoot, 1996) and in the process develop quite different expectations for particular classes or individual students. These expectations reveal themselves in the form of teaching (with ‘weaker’ students being allowed to achieve a great deal less work than ‘better’ students) and the content presented (teachers do not feel the need to be as well prepared with the ‘weaker’ students as with the ‘better’ ones) that can lead to the realisation of the self-fulfilling prophecy in students (Rowntree, 1987). This helps ‘highs’ to become higher and ‘lows’ to become lower (Rowntree, 1987). Moreover when, as in traditional assessment, only one kind of ability is rewarded – that which ‘weak’ students, by definition, do not have – most students become demotivated and are effectively pushed out of the system (Broadfoot, 1996).

3.1 The Alternative Assessment Paradigm

Our new conceptions of learning (see introduction to the Literature Review) and the ‘limitations’ for the teaching-learning process highlighted in section 3.0 make it, in my view, unsustainable to continue using an assessment model that has traditionally developed to focus on selection, certification and accountability. In particular, the recognition that assessment is now required to achieve a wider range of purposes than this – which include supporting teaching and learning, providing information about students, teachers and schools, and driving the curriculum and teaching – has increasingly rendered the traditional model underpinning assessment theory an inadequate framework and has necessitated the development of a new theory to further our understandings of and practices in educational assessment (Gipps, 1994). It is to this new reconceptualisation of assessment, which I find so educationally promising, that I now turn my attention.

3.1.1 Philosophical Underpinnings of the Alternative Assessment Paradigm

When still in its infancy, Murphy and Torrance (1988) crystallised the concerns and hopes of the alternative assessment paradigm:

Our own predominant concern is not for assessments to be psychometrically pure and reliable, but for them to play a constructive role in the educational process and as a part of that role to provide valid information about educational achievement in the fullest sense of the meaning of that phrase. Such achievement is not that which is *easily measurable*, but that which is *desirable* in terms of the broad aims of those concerned with what children gain from the process of education. (p. 100) (emphasis in original)

Their comments heralded a new emphasis on assessment procedures that provide a useful picture of what students know and can do. It was a way of saying that assessment should reward holistic learning and encourage qualitative understanding (see Hildebrand, 1996). From this perspective, assessment “has a constructive focus where the aim is to help rather than sentence the individual; thus it emphasizes the individual’s achievement relative to him or herself rather than to others, or in relation to defined criteria” (Gipps & Murphy, 1994, p. 261). Implied within this new understanding lies the notion that the purpose of assessment has changed from

categorising students for assignment to pre-determined curricular and instructional programmes to tailoring instructional programmes to learners' individual needs and to connecting learners and groups of learners in mutually beneficial learning experiences (LaCelle-Peterson, 2000). This necessitates a focus of attention on each student's learning, as everyone is believed to have his or her own particular requirements. Rather than unrealistically seeking equality of outcomes and the provision of identical experiences for all, equity in assessment is thus about assessment practices and interpretation of results that are fair and just for everyone (Gipps & Murphy, 1994). The embedded emphasis on the realisation of everyone's potential reflects in turn the current repositioning of assessment in relation to learning – that is, a change from being a 'measure of learning' to becoming a 'support to learning'.

Assessment in the new paradigm is in fact no longer seen as a scientific or objective activity. In recognition of the postmodern notions of multiple realities, subjectivity and knowledge construction, it is perceived instead as “an inexact matter and can never be an exact one” (Harlen, 1994b, p. 139). This shift from the psychometric model of assessment to the educational model – just like the one between the experimental and the naturalistic paradigms (see section 1.1.2) – draws on the postmodern condition that requires a suspension of belief in the absolute status of 'scientific' knowledge (see Gipps, 1993). This means that no matter how much we try to calibrate the 'measuring' instrument, we can still never know what is inside a student's head. Assessment can instead only tell us what the student can do in particular circumstances. Keeping in mind that domains and constructs are multi-dimensional and complex, that assessment is not an exact science, and that the interaction of student, task and context is sufficiently complex, we still cannot know what the student can do in other circumstances. This renders the generalisability of assessment to other tasks and contexts limited, if not dubious. As a consequence of the decline of external normative constraints, there is now more need for individuals to discipline themselves through internal mechanisms and a corresponding search for ways in which this may be achieved (Winch & Gingell, 1999). This calls for what Ellis (2001) terms 'reflective assessment' that requires practitioners, including students, to engage in a metacognitive assessment process in order to foster “a practical sense of what works, what is good, what has meaning, and ultimately, why” (p. xv).

The embedded denial within the new paradigm of the existence of a ‘true score’, however, reconceptualises rather than bans the use of tests and examinations in assessment. Constructivist theories demand in fact that tests show what students know and can do, as well as facilitate good learning – what Glaser (1990) calls ‘placing tests in the service of learning’. Within this emerging framework, tests should consequently be “ambitious instruments aimed at detecting what mental representations students hold of important ideas and what facility students have in bringing these understandings to bear in solving their problems” (Shepard, 1991, p. 9). The retention of tests evidently builds on the fundamental notion that “assessment does not stand outside teaching and learning, but stands in dynamic interaction with it” (Gipps, 1994, p. 15). It is all about fostering a much more comprehensive system that supports the appropriate use of multiple methods of assessment that are needed to cover the full range of achievements targeted – namely, knowledge, thinking, processes, products and dispositions (Gipps, 1994). The understanding is that the multiplicity and variety of assessment practices provides higher quality (as the strengths in one source compensate for the weaknesses in others) and fairer information (as all assessment methods may be said to have a certain amount of bias). This is in line with the growing realisation that “assessment is more than just a technical activity; it is a human activity that influences and affects many people” (Airasian, 2000, p. 22).

With students being the most vulnerable in this respect, it becomes imperative to always strive to obtain trustworthy information before making important decisions that can influence them. Moreover, once the information is collected, there is a responsibility to protect its privacy, to recognise its decision-making limitations, and never to use it to demean or ridicule a student. These ethical standards form an integral part of the new paradigm that, as Broadfoot (1996) points out, is driven above all by the desire to channel the powerful impact of assessment to promote, rather than inhibit, learning. This understanding explains why quality in assessment now depends mostly on the impact of assessment on learning rather than the accuracy of assessment *per se* (Broadfoot, 1996). The emerging idea that the indicator of the quality of educational provision is the individual students’ learning outcomes has now made it necessary to begin considering the traditional notions of reliability and validity in relation to the contexts and purposes of assessment. Thus, for instance, within the classroom, where

the teacher wants information about students on a regular basis with minimum interruption of normal work,

... quality in assessment means an assessment made and interpreted on the spot which provides the type of information required (high validity) and with the greatest degree of reliability possible in the circumstances. (Harlen, 1994a, p. 13)

This distancing from the traditional conceptions of reliability and validity is again evident in Gipps' (1994) understanding of quality assessment, which she defines in terms of 'trustworthiness'. The qualities of a trustworthy assessment include: (i) *credibility* – which comes from regular ongoing assessment in the classroom, and the inclusion of parents in the assessment dialogue; (ii) *transferability* – which requires the assessor to specify the context in which a particular achievement is demonstrated, so that others may judge whether this is transferable to other contexts; (iii) *dependability* – which makes the assessment process open to scrutiny, and subject to an audit process of quality control; and (iv) *authenticity* – which depends on the extent to which the relevant constructs are fairly and adequately covered in the assessment.

3.1.2 Mathematics and the Alternative Assessment Paradigm

The reconceptualisation of assessment that I outlined in section 3.1.1, which raised a number of issues that apply to all disciplines, subsequently stirred debates that have set forth fundamental conditions that form a foundation on which to build new approaches to assessment within the different disciplines, including mathematics (which is the field of practice in this study). One of the more influential 'mathematics initiatives' in this respect has been the publication of the American document *Measuring what Counts* (Mathematical Sciences Education Board [MSEB], 1993) that laid out a conceptual framework aimed at shifting assessment in mathematics away from simply measuring results to also contribute to the educational process itself. This framework – which was based on the three key principles of 'content', 'learning' and 'equity' – made the case that assessment in this domain should primarily reflect important mathematics, support good instructional practice, and enhance every student's opportunity to learn. Later on, NCTM (1995) used the MSEB document as groundwork for its detailed assessment standards. I find that the six criteria formulated by NCTM (1995), which are research-based yet written in a style that is easily

accessible to practitioners, are foremost amongst the number of guiding principles that have been prepared to facilitate the transition in mathematics assessment from the tradition psychometric paradigm to the alternative one. According to NCTM (1995) – and I cannot but agree – these statements are essentially about what is valued in mathematics, and reflect in their totality a vision of exemplary mathematics assessment (see Table 3.1).

Table 3.1: NCTM assessment standards for school mathematics

1. **The mathematics standard:** Assessment should reflect the mathematics that all students need to know and be able to do. Thus, assessments are based on significant and correct mathematics, and assessment activities provide all students with opportunities to formulate problems, reason mathematically, make connections amongst mathematical ideas, and communicate about mathematics.
2. **The learning standard:** Assessment should enhance mathematics learning. Although assessment is done for a variety of reasons, its main goal is to advance students' learning and inform the teachers as they make instructional decisions.
3. **The equity standard:** Assessment should promote equity. Equitable practices honour each student's unique qualities and experiences, and set high expectations for all. The goal is no longer now to have all students learn some mathematics, but rather for all students to develop their mathematical power to the fullest.
4. **The openness standard:** Assessment should be an open process. Open assessment, which involves shared responsibilities by students, teachers, and the public, has three facets: (i) making the information about the process available to those affected by it; (ii) honouring the professional involvement of teachers; and (iii) being open to scrutiny and modification.
5. **The inferences standard:** Assessment should promote valid inferences about mathematics learning. An inference about learning is a conclusion about a student's cognitive processes that cannot be observed directly. It follows that the validity of teachers' inferences depends on their expertise and the quality of the assessment evidence they have gathered.
6. **The coherence standard:** Assessment should be a coherent process. This involves three types of agreement: (i) the different phases of the assessment process fit together to form a coherent whole; (ii) the assessment matches the purposes for which it is being done; and (iii) the assessment is aligned with the curriculum and with instruction.

These assessment standards form part of NCTM's vision of reform in the content, teaching, and learning of school mathematics that is based on the assumption that 'all students are capable of learning mathematics'. NCTM (1995) uses the term 'mathematical power' to capture the shift in expectations for all students – which is at the heart of the alternative assessment paradigm – towards understanding concepts and skills; drawing on mathematical concepts and skills when confronted with both routine

and nonroutine problems; communicating effectively about the strategies, reasoning, and results of mathematical investigations; and becoming confident in using mathematics to make sense of real-life situations. The term is used in recognition that mathematics is more than a static collection of discrete concepts and skills to be mastered. Doing mathematics indeed includes such dynamic and integrative activities as discovering, exploring, conjecturing, sense making and proving. Mathematics students consequently “deserve a curriculum that develops their ‘mathematical power’ and an assessment system that enables them to show it” (NCTM, 1995, p. 11).

The understanding that mathematics is a dynamic set of interconnected, humanly constructed ideas calls for an assessment system that allows students to engage in rich activities that include problem solving, reasoning, communications and making connections (Romberg & Wilson, 1995). In particular, I would argue like Shafer and Romberg (1999) that unless the assessment programme is to be limited to checking whether students have acquired mathematical facts and procedures emphasised during instruction, it needs to create numerous opportunities for informal and formal assessment that involve mathematical reasoning at increasingly complex levels and that document growth over time. The ultimate aim is therefore for mathematics assessment to stretch over all the three levels identified by de Lange (1995). These are: (i) *lower level assessment* that deals with traditional mathematics and traditional tests, and concerns objects, definitions, technical skills and standard algorithms; (ii) *middle level assessment* that can be characterised by having students relate two or more concepts or procedures as they engage in making connections, integration and problem solving; and (iii) *higher level assessment* that deals with more complex material, such as mathematical thinking and reasoning, communication, critical attitude, interpretation, reflection, creativity, generalisation and mathematising.

3.2 Classroom Assessment

Assessment in education, which is the generic research area of this study, spreads over a number of closely-knit scenarios that vary primarily across a continuum of formality:

At one end of a dimension of formality, the task may be normal classroom work and the process of gathering information would be the teacher reading a pupil’s

work or listening to what he or she has to say. At the other end of the dimension of formality, the task may be a written, timed examination which is read and marked according to certain rules and regulations. (Harlen et al., 1992, p. 217)

However, given that this study is basically about Teachers' Classroom Assessment Practices (TCAP), the scenario that is directly relevant here concerns assessment that – irrespective of the level of formality, the assessors, and the type of tasks involved – originates inside the classroom as opposed to outside it. In the coming sections, apart from exploring this specific form of assessment, which I am calling 'classroom assessment', I also examine the implications that the alternative assessment paradigm bears upon it in order to be able then to better understand how the present assessment situation inside classrooms compares against the criteria of the new paradigm.

3.2.1 Defining Classroom Assessment

I like to think of classroom assessment – which I see as an umbrella term that transcends, albeit with changed emphasises, the different assessment paradigms – as the taking of the temperature of the teaching-learning environment inside the classroom. For me, it is the link between the assessment process and the place from which it originates (i.e., the classroom) that defines the term rather than 'the why, when and how the temperature is taken' or 'who takes and interprets the temperature to then decide on how to act'. It follows that although I am using a term that is much used in the American literature (e.g., Airasian, 2000; Anderson, 2003; Bright & Joyner, 1998; Brookhart, 1999; Bryant & Driscoll, 1998; Phye, 1997a,b; Popham, 1999), my understanding of it is slightly, albeit significantly I believe, different from the manner in which it is generally portrayed, even if only by implication. The recognition of student agency in learning, which has been a key element in the development of the new assessment paradigm, makes me critical of the numerous attempts in the literature (e.g., Anderson, 2003; Webb, 1998) to portray, possibly unwittingly, the teacher as the point of reference in classroom assessment at the almost complete exclusion of students. In my view, to accept this would be like saying that classroom assessment is the same as teacher assessment. Whilst this might have been possible to sustain within the traditional paradigm, my understanding that teacher assessment is "Assessment made by teachers of pupils' attainment, knowledge and understanding" (Gipps, 1994, p. 123) encourages me to argue instead that although teacher assessment is undeniably

an essential component of classroom assessment, it is not comprehensive enough to capture all the learning possibilities within the classroom offered by the alternative assessment paradigm. Indeed, according to this paradigm, it is only when students engage in self- and peer assessment (see section 3.2.3) – which I see as complementing teacher assessment inside the classroom – that assessment can be truly said to support, as opposed to inhibit, their learning.

Again, even though the rationales of teacher assessment and assessment by students are both linked to the constructivist model of learning (see Gipps, 1994) – and are meant as such to reflect primarily the learning needs of students – classroom assessment, which in my understanding incorporates both forms of assessment, is still not the same as formative assessment. Indeed, assessment is formative when it is “used to identify what pupils have learned, what they have not learned and where they are having difficulty. In this way, it supports the teaching-learning process” (Gipps & Murphy, 1994, p. 260). Formative assessment is moreover a low-stakes assessment situation that carries few, if any, long-term consequences for the teacher and students. It follows that whereas formative assessment refers to the function of assessment (i.e., the use of feedback into the teaching-learning process), teacher assessment and student assessment refer to the person who makes the assessment (i.e., the teacher and students respectively) (Gipps, 1994). This means that the classroom assessment process is formative only as long as teachers and students use it formatively.

It has however to be said that with teachers, the decision to use classroom assessment formatively or otherwise is often not just a matter of personal choice. For although teachers are in a position to continually gather unique information about students that offers great promise for formative assessment (Calfee & Masuda, 1997; Nuttall, 1993), they are often required to periodically collate and report assessment results to third parties. Indeed, as is mostly the case in Malta, even when their assessments are not used for external purposes, the school itself is likely to want them to generate assessment information for internal purposes. The possible subsequent use of teacher-generated information for both managerial (i.e., to select and certify students) and accountability purposes (i.e., to evaluate teachers, schools or age-groups at national level) (Gipps & Murphy, 1994) gives classroom assessment a summative dimension that, as Harlen et al. (1992) point out, is primarily concerned with summarising

information about student achievements at particular times. It is thus mostly because assessments made by teachers are not synonymous with formative assessment (see Torrance & Pryor, 1998) that classroom assessment can also be a high-stakes assessment situation that carries long-term consequences for students, and possibly the teacher and school.

3.2.2 The Purposes of Classroom Assessment

The literature frequently refers to the purposes and uses of assessment (e.g., Airasian, 2000; Black, 1998; Eisner, 1993; Gipps & Stobart, 1993; Murphy & Torrance, 1988, 1990; Rowntree, 1987). This is usually in a form of a list with an accompanying warning that the list in question is by no means exhaustive. Gipps and Stobart (1993) provide one of the more comprehensive lists. They identify six uses of assessment:

1. Screening: This refers to the process of testing groups of students, normally at primary level, to identify individuals who are in need of special help;
2. Diagnosis: This involves the use of tests to identify children's strengths and (more usually) weaknesses;
3. Record-keeping: Test scores and teacher assessments are put into student records to then help in the transfer process from one school level to the next;
4. Feedback: Results provide feedback about the progress of individual students and the teacher's success. On the other hand, results of classes can provide information to the school administration about the progress and success across the school, and school results can be used by outsiders to 'evaluate' schools and teachers;
5. Certification: A student is provided with a qualification that signifies that he or she has reached a certain level of competence or knowledge;
6. Selection: Students are selected into different institutions for further and higher education. They can also be allocated to different streams or sets within institutions.

I consider the above list to be fairly representative of the literature except for its lack of reference to the two closely interrelated functions of 'motivation' and 'control'. With motivation the idea is to use assessment in order to encourage student to learn. Thus, for instance, teachers and parents can and do appeal to the value of examinations in the job market as an incentive for students to behave well and work hard at school (Gipps & Stobart, 1993). But in a system that is based on rewards and punishment, motivating students to work can actually result in controlling them, in getting students to do something they might not otherwise be inclined to do. As the line between coercion and encouragement is indeed fine, I would argue like Rowntree (1987) that much therefore depends on the intentions and perceptions of the teacher and student, and the relationship between them.

Although Gipps and Stobart's list and my two additions, as is usually the case with other lists, refer to assessment in general, I find that all the included functions apply equally well to classroom assessment as defined in this study. This is particularly so in view of the close coexistence in classroom assessment between the formative and summative dimensions – a coexistence that is respected and reflected in the above enlisted functions. Gipps and Stobart (1993) themselves acknowledge that their six uses can be classified under 'professional' or 'managerial' nomenclatures according to whether the assessment helps respectively to enhance the educational process (read 'formative') or to manage and monitor the education system (read 'summative'). Their emphasis on the 'aftermath of assessment' is moreover in line with my understanding that, rather than the actual assessments, it is the interpretation of the resulting data that may distinguish formative from summative practice (William & Black, 1996). I do not thus consider the terms 'formative' and 'summative' to typify assessments, but as descriptions of the use to which assessment information is put (William, 1998).

Given that the same assessment can be used both formatively and summatively – even if, in general, an assessment would have been designed so as to emphasise one of these two functions (William, 2001) – makes it necessary to prioritise amongst the various purposes of assessment according to one's operating context. My understanding that classroom assessment should be more focussed on using the inferences drawn from the emerging information to inform instruction and to monitor day-to-day progress (Joyner, 1998) – which is to say that it should primarily support the learning process –

aligns my position with that of Gipps and Murphy (1994) who argue that the main purpose of assessment within the classroom context should be for professional rather than managerial or accountability purposes. Embedded in this vision lies the notion that, contrary to what invariably happens in the traditional paradigm, assessment should be used to maximise learning services for all students rather than to legitimise minimal learning services for many of them (see LaCelle-Peterson, 2000). The recognition that formative assessment is the type of assessment that really matters inside the classroom (Nuttall, 1993) builds on the realisation that certification and selection are artefacts of our social and educational system and that they, as such, are not central to the teaching and learning processes (see Gipps & Stobart, 1993).

Although my position is consequently that teachers have the duty as educators to focus on the formative function of assessment, I think it would be unyielding for teachers to drop their summative role for as long as educational systems, such as in Malta (see section 2.6.1), continue to be dominated by assessment for selection and certification purposes. For the summative dimension of teacher assessment can benefit the quality of such assessments (which are likely to continue for the foreseeable future) through the inclusion of skills, competencies and knowledge that cannot be assessed by the more traditional paper-and-pencil approach (see Broadfoot, 1996; Broadfoot & Black, 2004; also NCTM, 2000). But as far as the classroom reality goes, this improvement in assessment would have been of little concern were it not for the fact that assessment for selection and certification purposes supports instruction most strongly when classroom work and teacher's judgements are both valued and included in it (see NCTM, 1995). I would consequently argue that insofar as it is best that teachers are involved in both formative and summative assessment, the problem remains that of determining how teachers should go about satisfying the different requirements of these two roles.

The literature offers two main views as to how this tension may be eased, if not resolved. Some commentators (e.g., Harlen et al., 1992) argue that formative and summative procedures and data must be kept separate, as the use for summative purposes of information collected with the intention to support learning would severely impair its formative role. My understanding that it is the use to which assessments are put that determines whether the procedure is formative or summative (see above)

however makes me refute this position in favour of those (e.g., Wiliam & Black, 1996) who argue that to have two separate sets of procedures would produce an intolerable burden on teachers and that, in any case, in such circumstances the summative would always overshadow the formative because of accountability demands. The awareness that “If formative assessment is to prosper, initiatives aimed at supporting a positive link between formative and summative work are sorely needed” (Broadfoot & Black, 2004, p. 17) leads me to argue that teachers need to establish what Black (1998) calls a ‘workable relationship’ between their two distinct roles. This approach, which surely calls for goodwill, skill and discernment on their part, can help to avert the strain in the relationship between the teacher and students that, as Gipps (1994) warns, can result when teacher assessment is used for summative purposes, as the teacher may then be seen by students as judge rather than facilitator.

3.2.3 Classroom Assessment in the Alternative Assessment Paradigm

The current paradigmatic shift away from the traditional assessment paradigm that privileges assessment for non-professional (or non-learning) purposes inherently refutes the continuation of the system in which assessment is seen purely in terms of its product, its results, and the use of which results may be put in managing or even driving school systems (see Torrance, 1995). Assessment in the alternative paradigm is seen in fact as a process almost wholly integrated with teaching and learning (Torrance, 1995). This repositioning – which is in line with the deep transformation in our conceptions of learning, of assessment and of what counts as achievement – requires teachers and students alike to distance themselves from what Ellis (2001) calls an ‘objectified sense of assessment’, that is, the traditional view of assessments as events or objects that stand apart from teaching and learning. I therefore find myself agreeing with Gipps (1994) when she argues that, according to our current understanding, the appropriate assessment model inside the classroom is one that is designed to support the teaching and learning of important skills and concepts at both basic and higher levels. This newly emerging culture – which calls for an alternative way of comprehending classroom assessment away from the traditional psychometric model – is now generally known, at least in the UK, as ‘assessment for learning’. Contrary to the traditional ‘assessment of learning’ for the purposes of grading and reporting that has its own well-established procedures (ARG, 1999), ‘assessment for

learning’ is “the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there” (ARG, 2002). This ‘for learning’ process is specifically guided by ten research-based principles developed by the ARG (2002) (see Table 3.2).

Table 3.2: The ten ‘assessment for learning’ principles

<ol style="list-style-type: none"> 1. Assessment for learning should be part of effective planning of teaching and learning: A teacher’s planning should provide opportunities for both learner and teacher to obtain and use information about progress towards learning goals. It also has to be flexible to respond to initial and emerging ideas and skills. Planning should include strategies to ensure that learners understand the goals they are pursuing and the criteria that will be applied in assessing their work. How learners will receive feedback, how they will take part in assessing their learning and how they will be helped to make further progress should also be planned. 2. Assessment for learning should focus on how students learn: The process of learning has to be in the minds of both learner and teacher when assessment is planned and when the evidence is interpreted. Learners should become as aware of the ‘how’ as they are of the ‘what’. 3. Assessment for learning should be recognised as central to classroom practice: Much of what teachers and learners do in classrooms can be described as assessment. That is, tasks and questions prompt learners to demonstrate their knowledge, understanding and skills. What learners say and do is then observed and interpreted, and judgements are made about how learning can be improved. These assessment processes are an essential part of everyday classroom practice and involve both teachers and learners in reflection, dialogue and decision making. 4. Assessment for learning should be regarded as a key professional skill for teachers: Teachers require the professional knowledge and skills to: plan for assessment; observe learning; analyse and interpret evidence of learning; give feedback to learners and support learners in self-assessment. Teacher should be supported in developing these skills through initial and continuing professional development. 5. Assessment for learning should be sensitive and constructive because any assessment has an emotional impact: Teachers should be aware of the impact that comments, marks and grades can have on learners’ confidence and enthusiasm and should be as constructive as possible in the feedback that they give. Comments that focus on the work rather than the person are more constructive for both learning and motivation. 6. Assessment should take account of the importance of learner motivation: Assessment that encourages learning fosters motivation by emphasising progress and achievement rather than failure. Comparison with others who have been more successful is unlikely to motivate learners. It can also lead to their withdrawing from the learning process in areas where they have been made to feel they are ‘no good’. Motivation can be preserved and enhanced by assessment methods which protect the learner’s autonomy, provide some choice and constructive feedback, and create opportunity for self-direction. 7. Assessment for learning should promote commitment to learning goals and a shared understanding of the criteria by which they are assessed: For effective learning to take place learners need to understand what it is they are trying to achieve – and want to achieve it. Understanding and commitment follows when learners have some part in deciding goals and identifying criteria for assessment progress. Communicating assessment criteria involves discussing them with learners using terms that they can understand, providing examples of how the criteria can be met in practice and engaging learners in peer and self-assessment. 8. Learners should receive constructive guidance about how to improve: Learners need information and guidance in order to plan the next steps in their learning. Teachers should: pinpoint the learner’s strengths and advise how to develop them; be clear and constructive about any weaknesses and how they might be addressed; provide opportunities for learners to improve upon their work. 9. Assessment for learning develops learners’ capacity for self-assessment so that they can become reflective and self-managing: Independent learners have the ability to seek out and gain new skills, new knowledge and new understandings. They are able to engage in self-reflection and to identify the next steps in their learning. Teachers should equip learners with the desire and the capacity to take charge of their learning through developing the skills of self-assessment. 10. Assessment for learning should recognise the full range of achievements of all learners: Assessment for learning should be used to enhance all learners’ opportunities to learn in all areas of educational activity. It should enable all learners to achieve their best and to have their efforts recognised.

Reproduced from ARG (2002)

The fact that the phrase ‘assessment for learning’ has come to refer to “any assessment for which the first priority is to serve the purpose of promoting students’ learning” (Black et al., 2003, p. 2) may suggest, wrongly in my view, that ‘assessment for learning’ and formative assessment are necessarily one and the same thing. I agree with Black et al. (2003) that it is best to distinguish, even if subtly, between these two terms. For whilst it is true that most definitions of formative assessment (e.g., Gipps & Murphy, 1994 – see section 3.2.1) essentially echo and respect the spirit of assessment for learning, some others do not. This is particularly so with definitions (e.g., Nitko, 2001) that move beyond the immediate classroom environment to include scenarios such as when the teacher revises lessons or learning materials by using information obtained from their previous use. The point is that if assessment information about learning is used, say, for recording purposes or for long-term curriculum improvement without helping the learning of the students currently involved, it might be formative for the teacher, but not for the students (Black et al., 2003). Assessment for learning, on the contrary, is primarily about addressing the needs of the present students.

The classroom assessment reform vision upheld by the current drive towards assessment for learning, which Gipps captured so well in her book *Beyond Testing* in 1994, inevitably calls for changes in the traditional assessment roles of teachers and students. For just as the teacher is no longer considered to be a transmitter of knowledge but a facilitator of student learning and the student is no longer considered to be a receiver but a constructor of knowledge, so too are they no longer expected to behave respectively in classroom assessment solely as the ‘one who checks’ and the ‘one being checked’. It is thus no longer acceptable to see classroom assessment, as has traditionally been the case, simply as the gathering of information about students by the teacher in order to aid his or her decision-making process. The new paradigm calls instead for classroom assessment to be seen as the gathering of information by both the teacher and students about their teaching-learning situation in order to help them in their decisions. This emerges clearly from Cross’ (1998) attempt to provide an indication of their new roles in the reconceptualised classroom assessment process:

Classroom assessment informs teachers how effectively they are teaching and students how effectively they are learning. Through classroom assessment,

teachers get continual feedback on whether and how well students are learning what teachers hope they are teaching. And students are required, through a variety of classroom assessment exercises, to monitor their learning, to reflect on it, and to take corrective action while there is still time left in the semester. (p. 6)

Apart from the emphasis on the link between assessment and learning, the other key element of the new classroom assessment scenario highlighted by Cross (1998) is the realisation that students cannot rely exclusively on assessments made by teachers. For even though these assessments may provide them with good quality formative feedback (see section 3.4.4.2), students still need to become self-monitoring learners if their learning is to improve (Sadler, 1989). The active involvement of students in their assessment process is linked to the constructivist view that “it is essential to grasp the goals of one’s work and compare them with one’s present understanding if learning is to be meaningful and permanent” (Black, 1999, p. 126). The understanding is that if students are to become effective learners, they need to progress in their knowledge of themselves as thinkers and learners, in their understanding of particular tasks, and in their strategic knowledge of how to go about the improvement of their own learning (Alexander et al., 1991). The recognition that students’ progress depends on their coming to understand their strengths and weaknesses, and how they may deal with them (Harlen & James, 1997) leads me to argue that the capacity of students to judge their own work is more than a bonus in good formative assessment. It is instead a crucial component for developing the complex understandings through reflective habits of mind (Black, 1998) that they need to possess in order to share with the teacher in the responsibility for the quality of learning.

But students cannot become the hoped for independent and autonomous learners unless they first absorb standards of quality (Sadler, 1989) and develop the capacity for self-assessment (Black, 1998; Stefani, 1998). The onus that students are in a position to assume responsibility over their own learning falls, in my view, on the teacher as part of his or her mentoring role demanded by the constructivist theories of learning. Instead of remaining a provider of assessment information *per se*, the teacher should therefore also guide and counsel students about the quality standards expected from them and how to reach these standards. This would require the teacher “to share expertise, work in collaboration with students over setting assessment criteria, and provide exemplars and models so that students develop more of an understanding of

the learning goals for different levels of attainment” (Stefani, 1998, p. 348). It is however important that the assessment criteria being used are either criterion- or self-referenced. This would allow each student to learn respectively if he or she can do a specific task or range of tasks (Gipps & Stobart, 1993) and about his or her progress over time, building in the process a personal profile (Black, 1998). By doing away inside the classroom with the traditional norm-referenced criteria that supposedly measure how much better or worse a student’s performance is in relation to that of other students, we would be contributing towards the creation of a classroom assessment environment that fosters self-esteem and motivation in students, as well as encourages cooperation amongst them – what I see as essential ‘for learning’ ingredients that have been much neglected under the traditional paradigm.

Needs to be said, however, that the new vision presented here for classroom assessment cannot happen unless the ‘old’ undemocratic relation inside the classroom between the powerful (i.e., teacher) and the powerless (i.e., students) is replaced by dialogue in an environment in which the teacher and students alike feel comfortable to reveal their thinking (including the limitations of what they know). It follows that teachers have to let go their traditional inalienable right to rule their classroom as they move towards classroom dynamics in which teachers and students form a more collaborative community of learners (Peterson & Stack, 1998) – an important characteristic of the new paradigm that Cross (1998) seems to have missed (see above). Only then can students become insiders rather than consumers of classroom assessment (see Sadler, 1989). But the ensuing assessment partnerships between teachers and students cannot be effective without also developing partnerships in the teaching and learning process (Stefani, 1998). This explains why I am particularly in favour of creating the so called ‘communities of shared practice’ in which the students become “participants in a learning curriculum where understandings concerning what they are doing, or, in relation to assessment, what is being done to them are shared” (Elwood & Klenowski, 2002, p. 246; see also Wiliam, 1998). In such classrooms, the students not only learn to reflect on their work and their learning, and make critical self-judgements, but they also critique the work of their peers and use productively the critique of others (NCTM, 1995). This helps them to become self-directed lifelong learners who can carry on when they are out of formal education and no longer have people to direct them where, when or what to learn (Cross, 1998) – a prospect that is

much in line with what local political and educational authorities are trying to achieve (see section 2.6.2).

3.2.4 Assessment for Learning inside the Classroom: Where do we Stand?

I argued in the previous section that students' learning is best served when classroom assessment is guided by the principles of assessment for learning. This implies that the quality of classroom assessment, at least as understood within the new paradigm, depends on the extent to which there is a strong commitment towards the formative use of assessment practices in favour of the students currently in class. Using this criterion, the current local classroom assessment situation described in sections 2.2.2 and 2.6.1 leaves much to be desired. A look at the literature however reveals that the situation elsewhere is not that different from Malta because, although assessment reform has now been a major educational issue for almost two decades in many countries, assessment practices inside these countries' classrooms have not changed that much. The overall picture – which attests to how hard it is to introduce effective formative assessment into classroom practice (see Wiliam & Black, 1996) – depicts the formative assessment scenario in many countries as generally one of weak practice (Black, 1998). The main weaknesses are:

- 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-emphasized and the learning function under-emphasized.
- There is a tendency to use a normative rather than a criterion approach, which emphasizes 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, 1998, p. 111)

What I find particularly striking is the fact that the assessment situation in countries such as the US and UK, which have been and still are at the forefront of the assessment reform efforts, is nonetheless gloomy. This is, for instance, how Delandshere (2001) describes the present US assessment reality:

... educational assessment practices still reflect for the most part the legacy of the past: the purposes are narrow and the methods used generate limited data. The assumptions and theories of learning are implicit; examinees submit to the process without active and equal participation (e.g., critique, reflection, self-reflection), and secrecy, reward, and punishment remain key concepts. (p. 130)

Delandshere reports in fact that ‘alternative’ assessment practices in the US are more the exception than the rule, and that they are not particularly encouraged or supported in most schools. On similar lines, in the UK, the broader conception of educational assessment that was stimulated by the curriculum reform movement of the 1960s did not take hold and continues not to in view of the public’s need to get down to brass tacks, to go back to the basics, to measure, to monitor, and to mandate (Eisner, 1993). It seems that these residual values – which clearly mirror what Broadfoot (1995) calls the ‘Zeitgeist of a previous age’, that is, the set of unquestioned assumptions concerning the merits of particular techniques in the traditional paradigm – continue to this day to act as a major stumbling block to assessment change. The implication is that classroom assessment practices cannot change in line with the demands of the new assessment paradigm unless we first undergo a major change in the way we think about assessment in particular and education in general. This emerges clearly, in my view, from our manifest inability to move from the generally agreed-upon policies to related practices – a reality that seems to justify Eisner’s (1993) assertion that the problem is not one of correct policy formation but of practice (see section 3.3.3.3).

3.3 Teachers as Classroom Assessors

Irrespective of what the national assessment policies try to achieve, it is the assessment orientation of the teacher him or herself that makes the difference at classroom level with regards to what forms and purposes of assessment are emphasised and promoted within (see section 3.3.3). This is even more so in educational systems, such as in Malta where only a few schools have as yet developed an assessment policy to guide their assessment practices (Grima & Chetcuti, 2003) and probably fewer still are actually monitoring these practices, in which the vast majority of teachers practically have a free hand on how to organise their classroom assessment environment. The decisive role played by teachers in determining the nature of classroom assessment means that although, as the ARG (2002) rightly points out, much of what teachers and students do in classrooms can be described as assessment, it is the assessment practices

of teachers that are by far the more influential component of classroom assessment (see section 3.2.1). For in the continued absence of a formal recognition that entrenches student assessment in the classroom assessment framework, the very existence of assessment by students remains either subject to teacher approval or else limited to its informal aspect characterised by students drawing insights into their learning from what they can ‘instinctively’ pick up from their classroom experiences.

This means that gaining insights into TCAP, which is the primary aim of this study, goes beyond shedding light on teachers’ assessment practices *per se* to offer an understanding into the wider, encompassing world of classroom assessment. This is particularly so in my study because although I speak of ‘practices’, my interest in TCAP – as evidenced by the research questions posed in section 1.3.1 – includes gaining understanding into teachers’ views on assessment, especially inside the classroom, identifying what teachers ‘do’ and ‘do not do’ and the reasons that motivate these assessment decisions, and exploring the implications that teachers’ assessment views and practices have for students’ learning. My quest for knowledge concerning TCAP is thus marked by a desire to understand the stance, position, feelings, experiences and views of teachers without loss of commitment to my own values that are linked to the notion of assessment for learning (see section 3.2.3). Whilst the embedded desire to construe knowledge and values from the multiple perspectives of persons, who unlike ‘objects of study’ have purposes and emotions, led to the qualitative case study methodology described in Part Three of the thesis, in the following sections I limit myself to discussing issues that are, in my view, particularly relevant to understanding classroom assessment by teachers. These include looking at what normally characterises teachers’ work reality, the level of teacher competence in classroom assessment, research-based models of classroom assessment by teachers, and classroom assessment by teachers in relation to ‘others’ and external assessment.

3.3.1 The World of Teachers

Classroom assessment practices by teachers do not exist in a vacuum. They are part instead of a much wider teacher reality that both encompasses and influences what teachers do inside their classrooms, assessment included. The decision to depict below, albeit briefly, this ‘world of teachers’ is based on my belief that one first needs to gain

an understanding of this embedding context before one can fully appreciate TCAP and what motivates them.

According to many teachers, initial teacher training provides theory that is not related to practice (Day, 1993). As a consequence of their view that educational theory is largely irrelevant, much of what teachers learn is on-the-job, in response to classroom actions and colleagues who are for most teachers the main influence during their working lives (Day, 1993). Seeing teaching as a craft and an art (Mitchell, 1999), many teachers proceed on impulse and intuition, relying on personal experience rather than on reflection and professional education (Fang, 1996). The situation is such that trainee teachers seek to move from 'novice' to 'expert' status by searching for those 'recipes' that the experienced teachers appear to employ quickly and effectively through their tacit understanding of situations. Once acquired, this 'recipe knowledge' tends to be treated as unquestioned and unquestionable. In fact, many teachers eventually get so locked into a view of themselves as technical experts, that they start finding nothing in their world of practice to occasion reflection (Schön, 1983). This probably explains why teachers' practice, as Calderhead (1987) points out, is difficult to change once it has become established, routine, and adapted to its context. Teachers' reluctance or inability to change (see Broadfoot, 1996) is aggravated further by the lack of support that they receive to become reflective practitioners (see Schön, 1983). For whilst the concepts of reflection and reflective practice have become mainstream in the academic and educational research community, professional contexts still do not encourage or support either reflective practitioners or reflective practice (Cole, 1997). In these circumstances, teachers usually come to base their decisions on a simple yet deeply influential 'sense of practicality' – they have a powerful sense of what works and what doesn't for 'this' teacher in 'this' context, as opposed to in the abstract or as a general rule (Hargreaves, 1994).

Many teachers hold a simplistic 'bright-person' model of teaching that is based on the delivery of information by the teacher and the decoding of that information by students (Sedlak, 1987; cited in Fang, 1996). This leads to a teaching-learning scenario that follows primarily the traditional model of 'explanation and imitation' (see Swan, 2001). Ironically, although teachers define their own success in terms of student learning (Airasian, 2000), few of them know much about the learning process itself

because most of them only have their past 'easy learning' experience to guide them (Cross, 1998). The learning situation, especially at the post-secondary level and beyond, is however different for many of their students. Nowadays, as more students continue studying beyond the compulsory years, many are finding learning at this level difficult and threatening (Cross, 1998). The teachers' general reaction to this is what Ellis (2001) calls the 'coverage game' – a concentrated effort on their part to make sure that all the course content is covered irrespective of the depth that is often lacking. This suggests that teachers' minds are not particularly attuned to promote learning. On the contrary, the difficult conditions in schools, such as large class sizes, unreasonable curricular and professional demands, lack of resources and supports, and numerous and persistent outside interferences, create high levels of anxiety for teachers and force many into a 'survival' mode (Cole, 1997). It remains however a precarious 'survival' exercise because teachers are answerable to a variety of 'clients' – that apart from students, whose education remains the main scope of teachers' activity, include others such as parents, administrators, advisers, inspectors, employers, curriculum development agencies, and politicians – whose expectations can be in conflict with each other as well as with the beliefs of the individual teacher (see Calderhead, 1987).

The physical and psychological structures of schools perpetuate norms of isolationism that render them lonely places for teachers to work in (Cole, 1997). In line with the negligible opportunities that teachers have at school to interact with colleagues in meaningful ways (Schoenfeld, 1999), 'individualism' is, according to Hargreaves (1994), the most prominent characteristic of teacher culture. This entails many teachers teaching alone, behind closed doors, in the insulated and isolated environment of their own classrooms. But whilst classroom isolation offers many teachers a welcome measure of privacy, a protection from outside interference that they often value, it greatly limits their access to adult feedback on their value, worth and competence (Hargreaves, 1994) and helps to engender lack of trust amongst teachers (Cole, 1997). With regards to control over their work, teachers still have little of it in spite of recent 'teacher empowerment' initiatives (Cole, 1997). Schools remain overly centralised institutions in which teachers have substantial influence over some issues of classroom instruction, but have little influence over schoolwide policy matters (Ingersoll, 1996). The traditional hierarchical and isolationist structure of schools and school systems have created contexts in which teachers are allowed to make operational decisions

inside their classrooms, but not to have a major voice where decisions that affect their work and development are taken. Teachers thus only have a measure of classroom autonomy that, as Hargreaves (1994) reports, they guard jealously. The combination of such autonomy and isolated practices leads to what Gipps et al. (1995) call ‘autonomous isolation’, which is the antithesis of rich professional dialogue.

3.3.2 Teacher Competence in Classroom Assessment

As discussed in section 3.2, although teachers have both a formative and a summative role to play inside the classroom, the understanding that the teaching-learning process is best served by a form of assessment that supports learning demands an emphasis on the formative dimension of teacher assessment. This means that whilst teachers have to somewhat produce defensible assessment results for outside classroom use (see section 3.3.5), they must not do this at the expense of providing “information to be used as feedback by teachers, and by their students in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged” (Black et al., 2003, p. 2). For this would mean that even though teachers, as Torrance and Pryor (1998) correctly point out, are right at the heart of formative assessment in view of their possibility to collect good quality data and to make best use of it in their feedback, they are still not fully exploiting their privileged position. With teachers having to nurture and promote the formative dimension of assessment without neglecting the summative one, it is not hard to see why Eggleston (1991) argues that classroom assessment is a sophisticated, difficult and demanding task on teachers. The complexity involved calls in fact for teacher assessment competence to be carefully addressed during periods that precede, accompany and follow professional experience by teachers (Eggleston, 1991) in such a way that it becomes ‘integral’ to their ideology and classroom practices (see Broadfoot, 1995). I see in such training the means to weld theory and practice into a reflective experience that extends teachers’ assessment craft knowledge, which is after all the knowledge on which teachers rely in the face of challenge and complexity (see Day, 2005).

The point is that teachers need to be assessment literate if they are to get the full benefit of classroom assessment (Gipps, 1994). This is, however, not just a matter of providing teachers with good assessment instruments – they also need help to develop

methods to interpret and respond in a formative way (Fuchs et al., 1991; cited in Black & Wiliam, 1998). In order to assess well, at least as understood within the new paradigm, teachers have to understand the constructs they are assessing, they have to know how to get at student's knowledge and understanding, and they have to know how to elicit the student's best performance (Gipps, 1994). And this within a framework of teachers knowing and understanding a great deal about the curriculum, about students' thinking, and about various ways to assess (Bright & Joyner, 1998). Effective classroom assessment moreover depends on teachers' understanding that learning is a sense-making process that students need to experience for themselves (see Bright & Joyner, 1998). Only then can the student component of classroom assessment capture the attention and recognition it deserves.

The question however remains whether teachers are being prepared and helped to shoulder their assessment responsibilities. Sadly, the conclusion reached by the US national survey on teacher assessment literacy that teachers are ill equipped to successfully assess their students (see Plake & Impara, 1997) is much indicative of what the literature has to say about the quality of teacher assessment competence. Research suggests in fact that although teachers spend as much as one-third to one-half of their professional time involved in assessment related activities, they receive virtually no training in classroom assessment (Stiggins, 1992). Not only do many teachers receive no formal assessment training, but most of those who have are also out of date in this rapidly changing field (Impara et al., 1993). This means that teachers are trained at best to know about grading, not assessment (Tolley, 1989). With most assessment decisions being as yet in the hands of teachers who still, for the most part, hold beliefs about learning and knowing grounded in early behaviourist principles (Shepard, 1991), it is little surprising that, as Mavrommatis (1997) points out, teachers tend to see assessment primarily in terms of selection, certification and accountability. With educational systems generally placing increased importance on grades as a measure of student progress and success as the school grade level increases (Stiggins & Bridgeford, 1985), teachers are practically being encouraged to believe that assessment and grades are inextricably linked (see Peterson & Stack, 1998). This probably explains the dominance of the grading function in classroom assessment (see Crooks, 1988). As a matter of fact, in line with the teachers' documented inability to understand well formative assessment (Black & Wiliam, 1998), their judgements are

often made on the basis of ranking student work rather than evaluating it against criteria (Gipps, 1994) and many teachers ‘assess to assess’, that is, they assess to assign grades rather than to learn about students (Bright & Joyner, 1998). In line with teachers’ tendency to give a summative meaning to many of their assessments (Black, 1998), marking – which often lacks frequency and quality (Office of Standards in Education [OFSTED], 1998) – is used to collect summative information and little attempt is made to use it formatively (Weeden et al., 2002).

Cizek et al.’s (1995) study suggests that although it is often claimed that teachers acquire much of what they know about assessment ‘on the job’ or through ‘trial and error’, the assessment practices of many teachers actually reflect their own individualistic values and beliefs. The many ‘filters’ that influence what/how teachers assess include the information that teachers decide is important to gather, teachers’ content knowledge, teachers’ views about the nature of the subject, teachers’ beliefs about the capability of their students, the nature of the curriculum being used, and teachers’ perceptions of community expectations (Bright & Joyner, 1998). In particular, the literature presents a teacher assessment reality characterised by strong parallels between their largely traditional views on teaching and learning (see section 3.3.1) and their assessment practices that are often still rooted in the psychometric tradition. The fundamental problem with this enduring reality is that classroom assessment is failing to inform teachers about students’ progress and is not helping students to learn about their learning. What I find particularly disturbing about this ‘not for learning’ situation (see section 3.2.3) is that the primary source of help with assessment that teachers have at school is their colleagues who have little or no training in assessment themselves (see Stiggins, 1988). Understandably, this inter-teacher dependence limits assessment knowledge and practices to those of current teachers (Bright & Joyner, 1998) – a reality that is certainly impeding the realisation of the new assessment paradigm into classroom practice. Whilst I find nothing intrinsically wrong with the long existing custom of teachers acquiring ‘competence’ in assessment by moving from novice to expert under the tutelage of a connoisseur (see Sadler, 1989), I am deeply concerned that especially in times of desired change, as at present, the mentoring teachers, albeit experienced craftsmen in their own right, are most likely to lack a proper understanding of the ‘why’ and ‘how’ of the intended

reform. Not only would this render them hardly suitable to lead others, but it also calls for their participation in retraining programmes.

Although assessment currently enjoys high centrality in educational discourse (see section 1.2.2), many teachers still cannot clearly describe the techniques, the processes and the beliefs they use as they form their assessment judgements (Bryant & Driscoll, 1998; Gipps et al., 1995). It is as if teachers consider assessment as a secondary activity. Suffices to say that teachers prefer to focus on teaching rather than on assessment activities (Airasian, 1991; cited in Mavrommatis, 1997) or learning (Black et al., 2003). And in their list of priorities, it is the setting of activities and content coverage that receive more emphasis, not assessment (Calfee & Masuda, 1997). Again, only a negligible few teachers are prepared to explore and use new assessment approaches (Stiggins & Bridgeford, 1985) that are seen to require extra time, which they claim not to have, for planning and implementation (Shepard, 1995). The fact that teachers can still recognise a fine performance when they see one (Sadler, 1989), which is a positive thing in itself, does not however redeem in my eyes the inherent learning limitations of the current assessment situation. For classroom assessment is not just about recognising the learning achieved, but also about creating the right climate to make learning happen. By failing to give teachers the tools and the chance to develop expertise in assessment (see Stiggins & Bridgeford, 1985), it is the whole teaching-learning process that pays a high price:

Assessment is commonly done haphazardly, without knowledgeable planning, implementation, or interpretation in classrooms at all levels of education, with negative consequences for students in terms of the content they learn, the way they learn it, and the quality of judgments made about their achievement. (Schafer, 1993, p. 123)

The research evidence indicates moreover that teachers hardly, if ever, have sound information about the progress of students' learning (Black, 1998). But teachers still manage to appear knowledgeable in view of their tendency to judge and describe students' performance in general terms rather than in terms of the specific assessment set (Airasian, 2000). It is therefore most sensible of Airasian (2000) to warn teachers to "be cautious about putting too much faith in their estimates of pupils' abilities and potential" (p. 210). I do however find it heartening that, as Impara et al. (1993) report, a number of studies show that teachers do not feel well prepared to perform their role

as classroom assessors. I see in teachers' declared concern and uneasiness about their assessment approaches (Stiggins, 1992; Stiggins & Bridgeford, 1985) a desire for improvement that can create the conditions for professional development. In the meantime, however, albeit the research evidence indicates that formative assessment by teachers can raise standards of learning (e.g., Black & Wiliam, 1998; Crooks, 1988), I must agree with Black (1998) that this potential is still not being exploited.

3.3.3 Models of Classroom Assessment by Teachers

The general lack of teacher assessment competence highlighted in the previous section does not translate itself in a uniform model of teacher assessment, not even within the same national context. On the contrary, there are indications that individual teacher values play an important role in determining the manner in which national assessment policies, which are now increasingly pressing for a closer link between learning and assessment, become classroom assessment practices. In the following sections, I refer to two British empirical studies – one by Torrance and Pryor (1998) and the other by Gipps et al. (1995) – that have identified different models of classroom assessment by teachers. Apart from briefly describing these models, I also explore what their side-by-side coexistence inside a system that is trying to align classroom assessment to the new assessment paradigm tells us about the prospects of change in assessment.

3.3.3.1 The Torrance and Pryor (1998) Study

Torrance and Pryor (1998) identify two conceptually distinct approaches to classroom assessment that they term 'convergent' and 'divergent'. These approaches, which are not necessarily mutually exclusive, seem to arise from teachers' differing views of learning and the supporting relationship between assessment and learning. Convergent assessment, which is rooted in a behaviourist view of learning, aims to discover whether the learner knows, understands or can do a predetermined thing in order to then teach or assess the next predetermined thing in a linear or at least pre-planned progression. The practical implications of convergent assessment, in which students are assessed by the teacher, are: (i) precise planning and an intention to stick to it; (ii) tick lists and can do statements; (iii) an analysis of the interaction of the learner and the curriculum from the point of view of the curriculum; (iv) closed and pseudo-open

questioning and tasks; (v) a focus on contrasting errors with correct answers; (vi) judgemental or quantitative evaluation; and (vii) involvement of the student as recipient of assessments. This view of assessment is more in line with repeated summative assessment or continuous assessment than with formative assessment.

Divergent assessment, in contrast to convergent assessment, emphasises the learner's understanding not the agenda of the assessor. The important thing here is to discover what the child knows, understands or can do. Upholding a constructivist view of learning, the intention is to teach in the zone of proximal development, and assessment is accomplished jointly by teacher and student. The practical implications of this approach are: (i) flexible and complex planning that incorporates alternatives; (ii) open forms of recording; (iii) an analysis of the interaction of learner and curriculum from the point of view of both the learner and the curriculum; (iv) open questioning and tasks; (v) a focus on gaining insights into current understanding and on prompting metacognition; (vi) descriptive rather than purely judgemental evaluation; and (vii) student involved as initiator as well as recipient of assessment. This view mirrors contemporary theories of learning and the complexity of formative assessment.

3.3.3.2 The Gipps et al. (1995) Study

Gipps et al. (1995) identify three models of teacher assessor within the implementation of the national curriculum assessment system in England and Wales – namely, the 'intuitives', 'evidence gatherers', and 'systematic planners'. Whilst these names serve to summarise each group's characteristics, the first and third models have within them two fairly distinct subgroups. The three models embody a range of approaches related to teachers' views on teaching and learning, and their general style of organisation.

For the 'intuitives', assessment is a kind of 'gut reaction'. Intuitives rely upon their memory on what students can do and use an all-round close knowledge of students built over time when making their assessments. 'Children's needs ideologists' and 'tried and tested practitioners' form two distinct subgroups within this broad group. Teachers in the former group, which holds a child-centred view of education, prefer a 'whole approach' to assessment as they see teaching and assessment taking place simultaneously. Their strategy is to record mentally all the time while watching the

processes a student is going through. On the other hand, teachers' close knowledge – which is often rooted in long careers of teaching – is the main basis on which the tried and tested practitioners make their assessments. These teachers have confidence in their self-grown assessment methods even if they do not always explain clearly the reasons for their choices. Their focus is moreover very much on teaching – the teacher first defines the student's needs and provides appropriate instruction, then the student responds, and finally the teacher marks and gives feedback. Assessment tends to be summative here, taking the form of work in exercise books and giving teacher-made tests or worksheets at the end of a lesson or topic. Believing that the actual class teacher needs to do the assessing, the tried and tested practitioners do not readily keep or consult previous records, and rely mainly on their own personal judgements.

'Evidence gatherers' are interested in collecting as much evidence (particularly written, as most of them prefer not to rely on memory) as they can, which they only evaluate later. Apart from sharing a basic belief in the primacy of teaching rather than assessing, evidence gatherers also believe that students generally learn what is taught and only what is taught. They thus see assessment as following teaching in order to check that the process is going according to plan. Evidence gatherers do not often plan assessment activities, but rely on the assessment 'opportunities' that arise within their normal classroom teaching. And when teachers plan assessment activities, these are primarily for the purpose of collecting written evidence rather than for diagnostic purposes. Their emphasis on written evidence arises from accountability concerns – the teachers want to have 'written proof' in case they are officially asked to justify their assessments. Typically, at the end of the term, all the available evidence is 'pulled together' so that the teacher can reflect on the work the student has done and 'weigh up' the overall performance. Their assessments consequently result from reflecting back over the student's performance in a summative manner.

'Systematic planners' plan for assessment on a systematic basis. A constructivist approach to learning underpins this model, with teachers expecting students to learn in idiosyncratic ways. Systematic planners are willing to let students get involved at higher levels not yet taught, using such occasions for diagnostic purposes. In view of their social constructivist beliefs, they attach importance to interacting and arriving at shared meanings with both students and colleagues. For them, assessment is a cyclical

process – a kind of diagnosis showing how the student is doing on the set tasks, and their role is to take notes (not trusting to rely solely on memory) and to plan accordingly for the next activity. Systematic planners also prefer to use many and varied assessment techniques, and are keen to fit the assessment technique to the activity and the student being assessed. Within the broad category of systematic planners, there are two identifiable subgroups – namely, the ‘systematic assessors’ who give daily concentrated time to assessing one group of students at a time, and the ‘systematic integrators’ who, instead of separating themselves off from the rest of the class for assessment periods, circulate to gather evidence in a variety of ways.

3.3.3.3 What do these Studies tell us about the Assessment Reform Process?

The coexistence of the above diverse models of classroom assessment by teachers within the same educational system (namely England and Wales) suggests that, in spite of mandated national assessment policies, individual teachers can still assess in a variety of ways, some of which can even be in conflict with the objectives of the intended policies. The lack of match between observed and desired teacher classroom assessment practices is, at times, so pronounced as to have convergent assessment practices (see section 3.3.3.1) and teachers following the ‘tried and tested practitioners’ and ‘evidence gatherers’ models (see section 3.3.3.2) – all of which are based on a traditional view of the teaching-learning process and its link to assessment – within a system that is actually trying to promote students’ learning as the principal aim of schools through an integrated, formative assessment process (see Task Group on Assessment and Testing [TGAT], 1988). The lesson to be learned here is, I believe, that no matter how well meant and/or how forcibly mandated new ‘for learning’ assessment policies may be, there can be no guarantee that these translate themselves into standard classroom assessment practices unless they are understood and desired by teachers. Policy makers would consequently be short-sighted to push on with the assessment reform before genuine efforts are made to align the practitioners’ values to the underlying values of the alternative assessment paradigm. This realisation is particularly relevant to the present study, embedded as it is within a Maltese context that has recently introduced alternative assessment policies in the hope – unfounded it would seem – of reforming its assessment practices (see section 2.6.2).

The difficulty of luring teachers away from their ‘traditional’ summative practices is well documented. In England and Wales, for instance, when teachers were told upon the introduction of the national curriculum assessment system in the 1990s to produce ‘Teacher Assessments’, many of them responded by setting summative and terminal tests that imitated the national assessment tests (McCallum et al., 1993; Torrance & Pryor, 1998). Likewise, in the Australian state of Queensland, the abolition of external tests in favour of sole reliance on school-based assessment produced little effect on classroom assessment practices at first – further state intervention was subsequently needed before finally detaching practice from reliance on stereotyped terminal tests (Butler, 1995; cited in Black, 1998). The undeniable ‘teacher hurdle’ to the current assessment reform efforts is however understandable. One only has to realise that the implementation of formative assessment, which is at the heart of the new assessment paradigm, calls for “rather deep changes both in teachers’ perceptions of their own role in relation to their students and in their classroom practice” (Black & Wiliam, 1998, p. 20). It is therefore not simply a question of swapping traditional practices for new ones – the needed reform touches in fact the very essence of what assessment by teachers should be all about. This emerges clearly from the five challenges that, according to Weeden et al. (2002), teachers face in developing assessment for learning:

- The *first challenge* is to clarify for themselves what they understand by ‘formative assessment’, and to decide how they can make initial (small) changes to their practice that will help students actively engage in their own learning, help them to be clear about their current performance and decide what they need to do next.
- The *second challenge* is to recognise that teacher expectations have an important effect on students’ learning and to look for strategies that have a positive impact on motivation and learning;
- The *third challenge* is to realise that although they may initially find formative assessment time-consuming and may appear to reduce the time for teaching, experience suggests that this issue disappears later.
- The *fourth challenge* is to learn how to collect and interpret the data formatively.

- The *fifth challenge* is to recognise that any change (which will need support from colleagues, parents and policy makers) does not happen overnight, and that they need to be able to experiment, share ideas and find out what works for them in their own context.

These challenges highlight the complexity involved in helping teachers to transform themselves from being a possible liability for change to becoming the means for change. With assessment practices reflecting and reinforcing the often conflicting values embodied in educational systems (Broadfoot, 1996), I am particularly convinced of the need to work on changing values, not just of teachers but also of society at large, if true assessment reform is to follow. At the same time, however, I realise that assessment reform cannot proceed on its own – it has instead to go together with reforms in curriculum and pedagogy (Black, 1998). For it is only within a holistic approach to change that improved assessment practices can, as the ARG (2002) argues, safeguard the necessary quality of learning experiences needed to achieve the goals of education. In other words, albeit improvement in assessment is necessary for raising standards in education, it is by no means a sufficient condition (Torrance, 1995).

3.3.4 Classroom Assessment by Teachers in relation to ‘Others’

In the traditional assessment paradigm, the teacher assumes almost total responsibility for assessing student learning (see Cross, 1998). Contrary to this reality of ‘isolated teacher practices’, classroom assessment assumes the characteristics of a wide ‘partnership’ within the new paradigm (see Stefani, 1998). In section 3.2.3, I referred to Stefani’s (1998) work in relation to one such partnership – namely, between teachers and students. But Stefani is, in actual fact, after a much more inclusive form of assessment partnership. Her argument is that assessment, at any level, should not be a unilateral activity, but should instead be carried out in partnership with interested individuals or bodies. This position, to which I adhere, heralds an important shift away from the ingrained secrecy that shrouds traditional assessment. The emphasis is now on an open form of assessment that is characterised by dialogue and collaboration at class, professional, and community levels (see NCTM, 1995; see also principles 4 and 5 of the American National Forum on Assessment – Phye, 1997b). This openness in assessment is basically about students sharing responsibilities with teachers for their

learning, the public (e.g., parents, policymakers, teacher educators, and business and industry leaders) being kept informed about the process, and teachers participating actively and working together in all its phases. The point however remains the extent to which, if at all, classroom assessment is actually following this openness route.

Sadly, the literature depicts a scenario still driven by the belief that assessment is part of the teacher's academic freedom and is thus his or her prerogative (see Brookhart, 1999). This is based on the 'misguided' understanding that only teachers can make 'qualitative judgements' that, as Sadler (1989) explains, are judgements that cannot be reduced to a formula that can be applied by a non-expert. As a matter of fact,

... many teachers perceive evaluation as the responsibility primarily of teachers because it constitutes part of the specialized knowledge and expertise that they have acquired as professionals. Assessment is regarded as strictly the teachers' prerogative: it sets them apart from their students and to some extent from parents and the rest of society. (Sadler, 1989, p. 141)

This probably explains why, as Cross (1998) points out, students are rarely involved either before or after providing the 'data for analysis'. And this to the extent that teachers often do not make their evaluative criteria or standards explicit to students – preferring instead to just imply them whilst keeping them inside their heads (Gipps, 1994). Given that their exposure to 'evaluative and editorial activity' is limited to as it is received from the teacher (Sadler, 1989), it is most likely that students' only involvement with assessment is their own self-directed version of 'assessment by students'. This occurs when during the ongoing interactions between teacher and student, and even between the students themselves, the students engage, often unconsciously, in an assessment of their performance in relation to each other as they strive to achieve the socially defined goals that underlie the teacher's activities (Broadfoot, 1996). This dependence on the teacher is in line with the fact that students' work is most often evaluated autonomously by the individual teacher (see Bryant & Driscoll, 1998). For there is little or no assessment collaboration amongst teachers (Cizek et al., 1995; Stiggins & Bridgeford, 1985), and teachers appear to be unaware of the assessment work of colleagues and do not trust or use their assessment results (Cizek et al., 1995; Hall et al., 1997). Assessment collaboration amongst teachers is apparently restricted to routine consultation with previous teachers in order to corroborate or reinforce current observations (Airasian, 2000).

The current classroom assessment situation, in which the public is invariably also kept at a distance, is thus a long cry from the desired partnerships that flourish in open environments. This ‘lack of involvement’, as I will shortly explain, carries serious repercussions. Having already discussed in section 3.2.3 the reasons why students ought to be involved in classroom assessment, I limit myself here to discussing the benefits of involving the public and fellow colleagues. To my mind, provided that the classroom assessment process follows the canons of the alternative paradigm, the greatest benefit of opening it to the public is the opportunity to show it for what it should be – that is, a concerted effort to help students with their learning. This may go a long way to help dispel the traditional, much-diffused notion that the primary function of classroom assessment is ‘to measure and report on student achievement’. Apart from likewise facilitating the acceptance of the new assessment philosophy, the specific involvement of parents harnesses moreover their enormous potential for facilitating students learning in more holistic ways (see Filer & Pollard, 2000).

There are, on the other hand, two main advantages when the classroom assessment process is open to participation by colleagues. The first, and I would say the more important one as it is directly linked to students’ learning, concerns the enhancement of the quality of teachers’ assessments. The very fact that teachers in schools with an open assessment process meet to discuss learning goals, expectations, students’ work, and criteria for evaluating achievement, helps them to form common definitions of what is to be assessed and to reach consensus on the appropriate evidence of students’ learning (NCTM, 1995). This makes it possible for teachers to build up mutual expertise that no one else possesses and that no one from outside can give them (Black, 1998). By participating in such ‘communities of practice’, teachers have moreover the opportunity to reorganise their knowledge of students, the curriculum, pedagogical practices, and possibly even the subject matter itself (Bright & Joyner, 1998). These communities – which lead to what Sadler (1989) calls ‘a form of guild knowledge’ – essentially serve to engage teachers in ‘reflective practice’ that entails “careful, critical examination and analysis of the underlying beliefs of our teaching and our practice for clarification and resolution of meaning” (Elwood & Klenowski, 2002, p. 254).

The second advantage, which is mostly relevant to the formal or high-stakes dimension of classroom assessment by teachers (see section 3.2.1), concerns the issue of

reliability. For especially, albeit not only (see Harlen, 1994a), when classroom assessment results operate in parallel with either school promotion procedures or external examinations, there must be some assurance to those receiving and using these results that there is comparability across teachers, tasks, and students (Gipps, 1994). Towards this end, I would argue after Pennycuick (1991) and Gipps (1994) that when teachers come together to talk about teaching and to reflect on what they are learning about students, they would do well to discuss examples of students' work in order to arrive at shared understandings of the criteria in operation both with regards to the processes and the products of assessment. These moderation procedures, apart from helping teachers achieve assessment results that give dependable information about students' performance, lead to an enhanced professional status and increased public confidence in their judgements (Harlen, 1994b).

3.3.5 Classroom Assessment by Teachers in relation to School and External Assessment

Assessment by teachers – which can be formative or summative according to the aftermath of assessment (see sections 3.2.1 & 3.2.2) – has both a 'formal' and an 'informal' component to it. As I use them here, formal and informal assessment parallel respectively the 'structured' (i.e., planned and systematically designed) and 'spontaneous' (i.e., arising spontaneously from the naturally occurring classroom environment) assessments described by Stiggins and Bridgeford (1985). A fundamental difference between the two is that whilst formal assessment (e.g., homework or class tests) provides permanent evidence of a more summative nature (e.g., scores or marks), informal assessment leads to momentary or ephemeral evidence that can be lost forever unless it is captured immediately (see Wiliam & Black, 1996). Their lack of 'hard evidence' – which is inconceivable within the traditional assessment paradigm – probably explains why teacher judgements based on informal assessment carry, undeservedly I would say, less weight in the system.

While these judgements are frequently based on sound experience, they are often not clearly articulated and thus underrated, labelled as subjective and accorded less status than numerically quantifiable results of written tests, although they usually reflect children's performance with greater subtlety. (Denvir, 1989, pp. 282-283)

But although the ephemeral evidence of informal assessment is traditionally shunned in the more formal assessment settings (William & Black, 1996), it has an undeniable appeal amongst teachers inside their everyday classroom settings. For it allows the teacher to ascertain there and then whether students have a sufficient understanding of a concept and readiness to extend it further, or whether more work – and perhaps even what kind of work – is needed to first ‘cement’ basic understanding and performance. The strong link between informal and formative assessment is such that, as Denvir (1989) maintains, although teachers sometimes carry out formative assessment deliberately, this is often done unconsciously through activities selected intuitively. To their credit, teachers seem to appreciate mostly assessment experiences that push forward the teaching-learning process. This emerges from the number of studies that document how they favour their intuitive, idiosyncratic assessment strategies over their more formal, definite ones that are seen to produce only ‘dead data’ for reporting purposes (Broadfoot, 1996). Alas, just like the system pays little attention to the myriad of informal assessment opportunities that occur naturally in every lesson – which include teachers listening to students, observing them, and making sense of what they say and do (NCTM, 1995) – so do teachers almost systematically fail to include them when reporting on students at an official level. Because they are public, have important consequences for students, and must often be defended, these ‘official’ classroom data are usually based upon formal evidence like tests, projects or reports (Airasian, 2000). Teachers’ reliance on the results of formal assessments for official reporting probably relates to their relatively easy public defensibility:

It is defensible to say to a pupil or parent I gave a C grade because when I compared your test scores, projects, and homework assignments in this marking period to my grading standards, you performed at a C level. It is not defensible to say I gave a C grade because I *had a strong sense* that you were not working as hard as you could and because I *have a negative general perception* of your daily class performance. (Airasian, 2000, p. 213) (emphasis in original)

Despite many teachers’ lukewarm endorsement of their own official assessments, these remain a central classroom activity. Airasian (2000) explains why:

Pupils, their parents, and the public at large consider them to be very important and take them quite seriously. The grading, placement, promotion, and other decisions that result from official assessments influence pupils’ lives both in and out of school. They are the public record of a pupil’s school accomplishments and are often the sole information a parent has of how a child is doing in school. (p. 96)

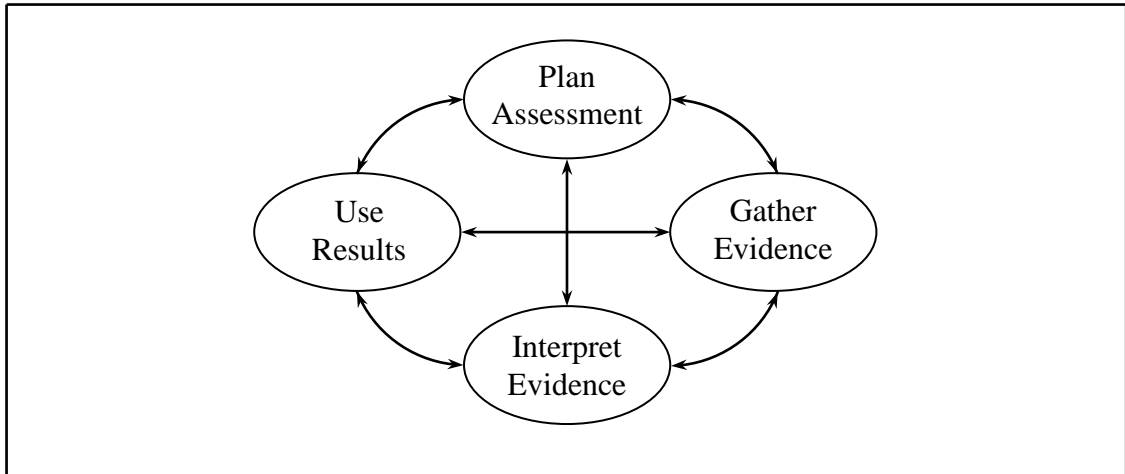
The resulting limited and biased form of information that is reaching people outside the classroom does not do justice to the possibly rich knowledge that teachers gain about students' learning through their informal assessment practices that, as Thomas (1990; cited in Gipps et al., 1995) contends, are probably one of the most difficult parts of their job. What I find, however, even more disturbing than this is that teachers, in their intent to prepare students for the selection and certification assessment procedures, apart from probably 'teaching to the test' (see section 3.0), allow their assessment practices to become inordinately influenced by these purely summative practices. In particular, as teachers feel obliged to help students succeed on external assessments, they are likely to feel pressured into making their own assessments 'look like', or at least similar, to external assessments – and, to achieve this, teachers familiarise themselves with the type of questions included in external assessments (Bright & Joyner, 1998). Albeit most teachers consider external examinations as one of the most fundamental constraints on their professional practice (Broadfoot, 1996), their main motivation is to provide students with preparatory test-taking experiences (Torrance & Pryor, 1998). Given that classroom assessment by teachers is supposedly primarily concerned with the formative function, I consider this mimicking of external summative assessments for classroom use as inappropriate for a number of reasons:

- In summative testing the need for a single overall result means that quite disparate data (for example, for practical and for theory) have to be added in ways that are often arbitrary: formative assessment does not have to do this.
- Summative assessment has peculiar problems with criterion referencing, partly because of the need to aggregate, partly because it cannot rely on personal judgments in deciding about the application of broad criteria to the work of individual pupils; such problems are far less serious in the practice of formative assessment.
- Summative work has to insist on standards of uniformity and reliability in collection and recording of data which are not needed in formative work and which inhibit the freedom and attention to individual needs that formative work requires.
- Whilst summative processes have to be seen to be 'fair' in treating all pupils in the same way, formative practice, with its priorities of identifying and helping to meet the learning needs of each pupil, can treat different pupils very differently.
- Summative tests are designed to give an overall score, so they work at low resolution and may yield little of the meaningful detail that formative feedback requires.
- Summative purposes can demand documented evidence for results – for example, for any auditing review – and so add to workload and distort formative practice, whereas formative work calls for action on the data rather than storage of it. (Black, 1998, pp. 120-121)

Ironically, there is evidence to suggest that teachers do not even trust the very same external assessment instruments that they so mimic, preferring instead to rely on the instruments that they themselves develop, such as teacher-made tests (Mavrommatis, 1997). Smith's (1991) study reveals in fact that the resulting numeric scores of external assessment mean little to teachers, particularly without the interpretive context (e.g., the vagaries of student effort and emotional status at the time of testing) that they alone possess. The literature, on the other hand, appears undecided on the extent to which teachers can actually predict the performance of their students on external examinations – the existence of which they only rarely question (Broadfoot, 1996). For instance, whilst two of the studies included in Black and Wiliam's (1998) review suggest that teachers can do it, Delap (1994, 1995) takes a more cautious position. He reports that, although there appears to be significant correlation between the predicted and the actual grades awarded, it is still prudent to treat teachers' estimates with caution because these estimates appear to be unreliable predictors of examination performance on an individual level. Again, whilst the predictive validity is somewhat higher when paper-and-pencil tests or examinations are used to predict the results of similar assessments (Nuttall, 1989), mathematics emerges as one of the subjects in which teacher predictions appear to be more accurate (Delap, 1995; Moody, 2001).

3.4 The Classroom Assessment Cycle within the Alternative Assessment Paradigm

Classroom assessment is normally presented as a cycle that is subdivided into a number of phases, often four (e.g., Bright & Joyner, 1998; Calfee & Masuda, 1997; Mavrommatis, 1997; NCTM, 1995). Using NCTM's (1995) terminology, these are 'plan the assessment', 'gather evidence', 'interpret the evidence', and 'use the results'. This division is however arbitrary because "In practice, the phases are interactive, and the distinctions between them are blurred. Assessment does not proceed through them in a neat, linear fashion" (NCTM, 1995, p. 4; see Figure 3.1). Instead, all classroom assessment episodes occur within a sequence of interrelated phases (Mavrommatis, 1997) to ideally form a coherent whole. In the coming sections, my main concern is to explore these four phases from the perspective of what the teacher can do in order to better promote the 'assessment for learning' philosophy that is now widely accepted to represent the role of classroom assessment within the new assessment paradigm.



Reproduced from NCTM (1995, p. 4)

Figure 3.1: The four phases of classroom assessment

3.4.1 Planning the Activity

Although, as argued in section 3.3.5, there is much to be gained for the teaching-learning process when teachers use formatively the information obtained from their ongoing informal assessment situations, one cannot simply let things ‘just happen’ for this would seriously jeopardise the quality of classroom assessment (see Bright & Joyner, 1998). In fact, given that classroom assessment, as it is now understood, is primarily about supporting learning, it is important for the teacher to gather as much ‘revealing’ information as possible, including through conscious planning, towards this end. The ultimate goal is for the teacher to have the necessary information to be able to plan work and to guide each student appropriately according to the learning goals of the course. Within this new paradigmatic scenario, even the transfer of assessment information between teachers assumes a ‘for learning’ as opposed to a ‘reporting’ dimension (see Black, 1998). Another ‘planning’ consideration that is linked to the new paradigm is teacher collaboration (see section 3.3.4). At issue is the need to substitute the traditional practice of teachers planning individually with the newer, open practice of teachers working as a group from the start, planning activities that would yield common assessment opportunities (see Torrance & Pryor, 1998).

Although I agree that teacher assessment plans should go well beyond the ‘next lesson’ context, I find Brookhart’s (1999) suggestion that these plans should be part of planning for a course from the very beginning a little unrealistic, especially in yearlong

courses. My choice to favour a more flexible approach to assessment planning does not however preclude me from arguing like her that teachers need to prepare thoughtfully and carefully their assessment activities – both at the overall and the day-to-day levels – by answering the questions ‘What kind of information is needed?’ and ‘What performance by students will give that information?’ This exercise, whose successful outcome largely depends on teachers having a clear sense of what they wish to assess and why they wish to assess it, should lay a solid foundation for the selection and use of proper assessment methods, which is an important prerequisite for quality assessment (see Stiggins, 1992). The emphasis in the new paradigm on linking assessment to learning does not however exclude that teachers, on certain occasions, other than using tasks for either diagnostic (before) or embedded (during) assessment purposes, also use tasks for mastery (final) assessment purposes (see Bryant & Driscoll, 1998). This is in line with the understanding that classroom assessment, apart from providing the information that is needed for immediate short-term purposes, is also used for summative long-term decisions (see Calfee & Masuda, 1997).

3.4.2 Gathering Evidence

In the new paradigm, the evidence-gathering phase is about gathering adequate and relevant information about students’ learning. The idea is to obtain as comprehensive a picture as possible about the teaching-learning situation. This calls for a gathering approach that, apart from tapping evidence from a variety of sources that are either pre-planned or that arise spontaneously during the lesson (see Airasian, 2000), is also guided by the aim of primarily seeking evidence that illuminates each student’s learning trajectory as opposed to comparing him or her against other students or norms. The evidence-gathering procedures that concern me here are directly related to the classroom situation – these include observation of students, written and oral communication, assessment tasks, and class tests. But the teacher may begin gathering information about students, both at individual and group levels, before even meeting them. For teachers are known to have their ‘antennae’ up at the start of school, constantly searching their environment for information about students – an exercise that Airasian (2000) calls ‘sizing up assessment’. Teachers obtain the outside classroom component of this ‘sizing up’ information from sources like the school grapevine, comments by other teachers, school records, and performance of siblings.

3.4.2.1 Observation

Much of the inside classroom component of this sizing up assessment comes from the informal observation of momentary unplanned happenings, such as a student doing or saying something, that the teacher mentally records and interprets (Airasian, 2000). The teacher uses this information, together with what is learned from outside classroom sources, to form an initial set of perceptions and expectations about students that will then influence the way in which he or she plans for, interacts with, and manages students and instruction (Airasian, 2000). My main concern with this reality is that, as Airasian (2000) points out, these early impressions – about whose accuracy teachers are generally very confident – tend to become permanent, virtually stable throughout the year. This means that the teacher often forms generalised and lasting impressions from early singular or limited instances – practically from what the student happens to be doing or saying when the teacher glances his or her way in the first few days or weeks. The unfortunate fact that assessment so used prejudices rather than aids learning leads in turn to what I consider as an even graver concern – namely, the stereotyping and labelling of students. For this triggers in students the mechanisms of the self-fulfilling prophecy that sees them going on to produce a reality that reflects these original evaluations by the teacher (see Filer, 2000).

Informal observation is however an important feature of classroom assessment right through the year, not only when teachers engage in their early sizing up assessment. Teachers do in fact rely heavily on observation to assess instruction (Airasian, 2000; Stiggins & Bridgeford, 1985). In particular, should their priority be to keep the teaching process going, they use the reaction of pupils to judge whether it is feasible to carry on (Black, 1998). Although, given the fast pace of classroom activity, it is quite understandable why, as Airasian (2000) contends, the primary indicators that teachers use to monitor instruction are those that are most readily available, most quickly surveyed, and least intrusive – basically, reactions from students such as facial expressions, posture, participation, questions, and attending – I would still argue like him that such reactions do not provide direct evidence of student learning, which is the real criterion of instructional success. What makes the situation even less tenable is the tendency of teachers, irrespective of whether this arises from seating arrangements or from their unconscious preference for certain students, to often focus on an overly

narrow sample of students and being inattentive to the rest (see Airasian, 2000). This facet of classroom assessment – which is characterised by the teacher’s probable impossibility to monitor the classroom experience of each student whilst fully engaged with instruction – is not conducive to the rich, individualised information that is needed to help students progress in their learning (see Calfee & Masuda, 1997).

This realisation does not however mean that there should be no place for informal observation in classroom assessment. For, in truth, even facial expressions may reveal relevant diagnostic information (Broadfoot, 1996). My position is consequently that since such observations produce evidence, however ephemeral, that has its own unique value, it makes good sense to continue with this evidence-gathering procedure. This can only be problematic ‘for learning’ should the teacher fail to collect also corroborative evidence through the use of multiple methods (see section 3.1.1). Other than this, the quality of observational evidence itself can be enriched should the teacher give, even if only at times, a more formal and planned dimension to his or her observations. Some studies show in fact that teachers find it surprisingly useful to suspend their active teaching – making clear to the class what they are doing and why – and to concentrate only on looking and listening with a few students at a time whilst the rest are engaged in individual or small-group activities (Black, 1998). This is in line with Calfee and Masuda’s (1997) assertion that assessment through observation improves when teachers create specific occasions for observation and practise ‘focus’ – that is, they select what and whom to observe, and put all else in the background.

3.4.2.2 Communication

Questioning is the most common form of evidence-gathering technique used by teachers when acting deliberately to obtain information about students’ knowledge or capabilities (William & Black, 1996). But albeit the teacher can question students both orally and in writing, it is the oral form that is normally practised. The popularity of oral questioning emerges from research showing that on average teachers may spend almost half their time on it (see Broadfoot, 1996). The almost parallel exclusion of written communication – which may take the form of the ‘pupil journals’ mentioned by Stiggins (1997) – works especially against teachers getting to know those students who, even when justly so are not publicly grilled through questioning, are ‘tongue-

tied' in classroom discussions and can express themselves better in writing. Whilst these missed 'written opportunities' somewhat undermine the rich data oriented spirit of the new paradigm, it must also be said that, as Brookhart (1999) points out, although questions in class help both teachers and students to clarify what students know and where their misconceptions have occurred, giving an accurate indication in the process of what the class as a whole understands, the information obtained from them does not give a complete picture of an individual's understandings. In truth, not only is there no guarantee that questioning elicits if a student has any particular knowledge or understanding, but one can also never be sure that a student does not know something through questioning (William & Black, 1996). I would consequently argue like Pryor and Torrance (2000) that teachers would do well not to treat the answers to their questions as unproblematic sources of information for pedagogic decision-making.

The need for teachers to reflect on the information received through questioning is also linked to what the different types of teacher questions – which are usually classified along a variety of interrelated categories – can possibly reveal about and contribute towards students' learning. With regards to categories, one can speak for instance of convergent and divergent questions and of higher and lower level questions (see Airasian, 2000). Convergent questions (also called 'closed') have a single correct answer and divergent questions (also called 'open') may have many appropriate answers. On the other hand, whilst lower level questions require students to simply retrieve and manipulate factual knowledge, higher level questions require students to build on this factual recall and engage in solving new problems. Although I agree with Airasian (2000) that there is a place during instruction for these different kinds of questions – suffices the knowledge that factual recall is the basis for higher level questioning – my position is that one needs however to prioritise amongst them. In particular, rather than emphasise closed questions that possibly leave students calculating whether to take the risk on their chances of knowing the right answer, teachers should invest more in open questions that show students that the teacher is interested in their ideas, encourage students' self-expression and challenge students to develop their thinking (see Black, 1998). Such a line of questioning that favours the creation of a classroom environment that is open to the potential of discussion as a learning tool (see Swan, 2001) parallels Torrance and Pryor's (1998) advocacy for the use of 'genuine questions'. Their argument is that unless questions elicit 'genuine' or

'authentic' responses from students (as opposed to prefabricated responses presumed by students to be what teachers want to hear), they would be grounded in the exigencies of teaching (i.e., to move the lesson forward – see also Broadfoot, 1996) rather than the promotion of learning. On the contrary, genuine questions, apart from providing insight into students' current state of understanding, are also potentially useful in stimulating further learning (see Torrance & Pryor, 1998).

But in contrast to this formative promise, Black and Wiliam (1998) conclude from their extensive review of the literature that the quality of classroom questioning is a matter of concern. Not only is questioning at all classroom levels dominated by recall questions (Stiggins et al., 1989) that follow the traditional ritual sequence of 'question by teacher, response by student(s), and feedback/evaluation by teacher', but teachers often also choose a sub-group of only a few students, and it is their reactions and responses to questions which serve to justify proceeding (Black, 1998). The overwhelming quantity of talk during classroom discussion moreover comes from the teacher, with very few words being actually spoken by students (Torrance & Pryor, 1998). This reflects a reality in which it is more common for a teacher's goal to be simply that of eliciting the correct answer from students rather than to engage them in discourse that requires them to articulate, develop and defend positions (see Calfee & Masuda, 1997). Teachers are so seduced into seeking and hearing correct answers, which then enables them to make a favourable judgement about their instruction, that they prefer factual questions to open-ended, complex ones in order to ensure more student participation and mastery (Airasian, 2000). This seduction is such that when they ask a question to students, they just move quickly around the class until they hear the right answer – conveying in the process an impression to students that it is speed that is important rather than thinking deeply about things (Boaler, 1997). The point is that questioning inhibits rather than helps the learning process when, as Lesh et al. (1992) contend, students are probably passed over by the teacher if they take more than three seconds to respond. Especially when faced with higher level questions, students need time to process their thinking in order to come up with more complete, thoughtful responses (Airasian, 2000). This calls for teachers to provide ample time and then to listen sensitively so that they pick up clues about a student's thinking that might need to be followed up (Black, 1998).

3.4.2.3 Tasks

For teachers not to rely unduly on the ephemeral evidence of classroom events, students need to systematically produce written work both in class and at home (Black, 1998). This calls, however, for assessment tasks that work towards valued learning goals and that are open in their structure to the generation and display of relevant evidence to the teacher and to the students themselves (Black & Wiliam, 1998). Towards this end, the teacher should select tasks that are “novel and varied in interest, offer reasonable challenge, help students develop short-term self-referenced goals, focus on meaningful aspects of learning and support the development and use of effective learning strategies” (Black & Wiliam, 1998, p. 31). In mathematics, for instance, this means that rather than going for tasks that students may enjoy doing but that contain little mathematics, or have little chance of developing mathematical understanding, or reveal little about students’ thinking (see Bright & Joyner, 1998), teachers should focus primarily on what NCTM (1995) calls ‘significant tasks’. These reflect current learning theories that configure the teacher’s role as that of helping students find, create and negotiate their meanings by providing them with meaningful and purposeful activities from their perspective (see Murphy, 1996). This positioning heavily curtails the use of atomised assessments – that is, when specific skills are assessed out-of-context rather than as part of a realistically complex task (see Black, 1998). At issue is the need for classroom tasks to be ‘authentic’ so as to facilitate the development of students’ understandings into knowledge that can be applied in real-life contexts, thus ensuring an explicit link between school learning and out-of-school practices (Murphy, 1996). These are tasks that, as Eisner (1993) points out, should:

- reflect the tasks that students will encounter both inside and outside schools;
- reveal how students go about solving a problem, not only the solutions they formulate;
- reflect the values of the intellectual community from which the tasks are derived;
- not be limited to solo performance;
- make possible more than one acceptable solution to a problem and more than one acceptable answer to a question;
- have curricular relevance, but not be limited to the curriculum as taught;
- require students to display a sensitivity to configurations or wholes, not simply to discrete elements;
- permit the student to select a form of representation to display what has been learned.

I see in this move away from the traditional decontextualised, rote-oriented tasks that impose low cognitive demands on students a shift towards a form of instruction that emphasises meaningful learning (see Darling-Hammond, 1994). Given that tasks constitute key contexts for students' thinking about the subject (Doyle, 1988), it follows moreover that the teacher, in his or her role of task selector, needs to possess a skilled and multi-dimensional foresight (Black, 1998). For not only must he or she reckon with constraints of time, of facilities, and of the starting-point of the students (Black, 1998), but attention must also be paid to the content of tasks as this sends a clear message to students about what parts of the subject are important to learn (Bryant & Driscoll, 1998) and the manner in which students are expected to work on the tasks as this delineates their learning habits. Our present understanding of the learning process makes it vital for students to be involved in collaborative projects, as these create the conditions for thinking aloud and sharing ideas, which is an important metacognitive aspect of learning and assessment so often lacking when students work alone in traditional school assessments (Ellis, 2001). Unfortunately, group work remains shunned by some teachers because they prefer to do all the talking themselves, by others because they prefer the silent atmosphere of a classroom where each student is busy doing his or her 'own' work, by others because they fear that this would limit the amount of work they can cover in a lesson (Ellis, 2001). And still by others because they have problems with student motivation, or what has been called 'free riding', which is a form of social loafing seen in a group when one or more members slack off and 'ride' on the extra efforts of their coworkers (Walker & Angelo, 1998).

Another important consideration in task selection that teachers have to grapple with is the degree to which the task is left open or closed. Whilst closed tasks are linked to standard textbook questions, school-learned methods and rules (i.e., tasks that encourage the development of procedural knowledge in students), open tasks are linked to practical and investigative work that requires students to make their own decisions, plan their own routes through tasks, choose methods, and apply their mathematical knowledge (i.e., tasks that encourage the development of conceptual knowledge) (Boaler, 1998). Moreover, whilst closed tasks typically have only one correct answer (i.e., close-ended) and a high element of scaffolding built in them (i.e., closed-middle, meaning that the task structure itself leads students towards a solution through known algorithms and procedures), open tasks are typically not tied to one

correct answer (open-ended) and are either less scaffolded or unscaffolded (open-middle). Apart from such considerations, a task can also be specified according to the complexity of reasoning it requires. Black and Wiliam (1998) refer in fact to a scheme developed by Dumas-Carre and Larcher in 1987 that can be used to produce such a comparative and descriptive analysis of tasks:

This scheme distinguished tasks which (a) presented a specific situation identical to the one studied, or (b) presented a ‘typical’ problem but not one identical to the one studied, requiring identification of the appropriate algorithm and its use, rather than exact replication of an earlier procedure as in (a), and (c) a quite new problem requiring new reasoning and construction of a new approach, deploying established knowledge in a new way. (pp. 31-32)

Clearly, as one moves from (a) to (c), the level of student thinking involved in working with tasks evolves from lower level (characterised by mere recall of factual information) to higher level (characterised by the application, analysis and synthesis of factual knowledge in order to solve new problems). But although teachers have such a wide array of tasks at their disposition – both with regards to openness and complexity – it is as if the thinking level demanded by a task is inversely proportional to its classroom use because, as Carter and Doyle (1987) point out, higher order tasks are rarely given in class. And when potentially demanding tasks are set, teachers avoid classroom conflicts by ‘redefining or simplifying task demands’ (Doyle, 1988). This teacher reluctance to spend time on what are basically nonroutine activities characterised by conceptual understanding, explorations, construction of meanings and invention – which is in direct conflict with the learning demands of the new paradigm – results from their perception (which is often correct) that these are irrelevant to students’ examinations (Goldin, 1992).

3.4.2.4 Class Tests

Testing remains synonymous with schooling (see Ellis, 2001). Not only do the majority of teachers spend more than 10% of their professional time on testing (Newman & Stallings, 1982; cited in Schafer, 1993), but teachers are also inclined to use tests irrespective of the purpose of assessment (Stiggins & Bridgeford, 1985). Teachers tend moreover to rely more heavily on paper-and-pencil tests when teaching mathematics and science than when teaching writing and speaking (Stiggins &

Bridgeford, 1985). One reason for this widespread use of written tests in mathematics is their ease of administration and collection of data in comparison to oral and practical assessments (Denvir, 1989). Needs however to be said that fair tests as such do not and cannot exist (Gipps & Murphy, 1994). It consequently makes no sense to expect tests to establish and to provide accurate feedback about what the student actually knows at a particular point in time (see Torrance & Pryor, 1998; also Gipps & Murphy, 1994). Apart from the test itself, its context may also make a significant difference. For instance, a low stakes test situation is unlikely to draw forth highly motivated best performances, which means that the data derived from such an exercise may not constitute a particularly valid indicator of educational achievement (Torrance, 1995).

In spite of its challenge to the traditional behaviourist rhetoric, the new assessment paradigm reconceptualises rather than abolishes the use of tests as evidence-gathering instruments (see section 3.1.1). The new emphasis on integrating tests to instruction in order to render them useful for instructional decisions (Black, 1998) builds on the understanding that it is the manner in which test results are interpreted and used by teachers and students alike that determines whether or not testing actually serves the formative or the summative function (see section 3.2.2). Notwithstanding their formative potential, I remain concerned with what lies behind the continued proliferation of testing inside the classroom. It is, for instance, worrying that whilst teachers are concerned about the time required to develop and use their own tests as this interferes with their instructional time, they tend to be less concerned about their lack of information on testing, their competence in testing, the student reaction to testing, and collaborating with others in testing (see Stiggins & Bridgeford, 1985). Probably of more concern is teachers' tendency to make little use of tests results beyond putting them into record books and using them to identify students for remedial help (Gipps et al., 1995). This little use of test results – which arises from teachers' inability to see tests as saying something about their teaching rather than just about the student (Wood, 1990) – is in line with teachers' general unwillingness to adapt the curriculum in response to testing (see Close & Brown, 1987; cited in Gipps et al., 1995). The 'unhealthy' distancing between testing and instruction is again evident in the manner in which class tests are designed to mimic the examinations used to certify achievement. This reproduces at classroom level the same problems that are generally associated with such examinations. In mathematics these include: (i) lack of

balance in the type of tasks presented; (ii) students required to give rapid responses in a limited time to numerous questions; and (iii) the message that the rapid use of well-learned techniques is most important (NCTM, 1995). The shortsightedness of this mimicry emerges also from the studies that show how class tests can improve student examination performance without any real or lasting improvement in educational quality (see Torrance, 1995; also Shepard, 2001 [cited in Schoenfeld, 2002]).

3.4.3 Interpreting Evidence

Collected evidence needs to be interpreted so that it may be turned into information on the basis of which decisions can be made (see Wiliam & Black, 1996; also Calfee & Masuda, 1997). As far as the teacher is concerned, the examination of the evidence helps him or her to determine whether or not there is a gap between what students can actually do and what he or she would like them to be able to do (Wiliam & Black, 1996). This reflective exercise helps

... teachers decide if instruction is being effective so that changes and modifications can be made. ... Some of the reflections will be formative 'evaluation' of students' progress, and some will be summative 'evaluation' that compares students' progress against established standards of performance. (Bright & Joyner, 1998, p. 31)

Although I hold that teachers should, given the present circumstances, retain their summative role (see section 3.2.2), I would still argue from an 'assessment for learning' perspective that the standards against which they compare the evidence should be primarily self-referenced or at least criterion-referenced, not norm-referenced (see Mavrommatis, 1997). This does not however exclude that the teacher, apart from interpreting the assessment data from a singular frame of reference in order to make decisions about single students, also views the data from a collective frame of reference in order to make group instructional decisions (see Phye, 1997a). In either case, for truly professional judgements, teachers need both time and occasion to think about the evidence, ideally in consultation with colleagues (Calfee & Masuda, 1997). For even if the interpretation of evidence is typically tacit and intuitive, based upon knowledge of students that teachers would have acquired through experience at both the collective and individual levels (Mavrommatis, 1997; Watson, 2001), it has to be said that teachers are so assailed by information in the classroom from all sides that

they rarely have the time to make considered decisions in the moment (Watson, 2001). It is again limiting on the quality of the interpretations made – as well as on students’ learning (see section 3.2.3) – that teachers are often, according to Wiliam and Black (1996), the sole interpreters inside the classroom of the assessment evidence.

3.4.4 Using Evidence

The interpretations that teachers give to assessment results are a means to an end. In fact, with very few exceptions, assessments are conducted for a purpose and certain actions follow the outcomes (Wiliam & Black, 1996). These actions or consequences, which are the focus of the next sections, can be grouped under three interrelated categories – namely, instructional decisions (which relate to the teacher interventions aimed at improving learning), feedback (which is mainly related to the formative function of assessment) and grading (which is primarily summative in nature). But given that these actions target different audiences, the need arises for teachers to have an adequate recording system in order to be able to select, edit and communicate assessment information appropriately and effectively. I find that Murphy and Torrance’s (1988) distinction between formative and summative recording provides the framework on which teachers can build such a system:

Formative records are essentially internal working documents, continuously updated and amended, for use by both teacher and pupil to encourage, guide and reward learning and to stimulate reviews of the curriculum and pedagogy by informing teachers of the effectiveness of teaching methods and the appropriateness of what is taught. Summative records are static, end-of-stage ... documents which present a distillation of all the assessment information available about a pupil geared, both in terms of content and format, to the needs or interests of audiences outside the school ... (p. 63)

3.4.4.1 Instructional Decisions

Teachers continually make instructional decisions according to their knowledge of what students know and can do (Bright & Joyner, 1998). Many of these decisions – which may involve proper instructional interventions or revision of tasks and assessments – are actually taken during the course of instruction itself on the basis of how teachers interpret ongoing assessment evidence (Airasian, 2000). In these circumstances, the teacher has to decide there and then whether the lesson is

progressing satisfactorily (in which case, the lesson continues according to plan) or whether a problem is sensed (in which case, the teacher either revises the planned instructional activity or initiates another teaching activity) – a teaching-assessment cycle that is repeated many times in the course of a single lesson (Airasian, 2000). The problem here is that it is hardly ever feasible for the teacher to monitor in detail the progress of each individual student during instruction. It is far more likely that the teacher monitors the impact of his or her teaching at an overall level rather than at the individual student level (Torrance & Pryor, 1998). Such a reality, characterised by the teacher focusing on individuals in detail only if they are causing real concern (Torrance & Pryor, 1998), unavoidable in practice as it may be, works against the realisation of each student's learning potential, which lies at the heart of the new paradigm. To act formatively, the teacher needs instead to have detailed, quality information about individual students (see section 3.4.2). For only then can he or she have the opportunity to put students in learning situations that are potentially optimal for them, and to optimise the activity and the learning process of each student within a given situation (see Perrenoud, 1998). It thus makes sense for the teacher to delay taking important instructional decisions, possibly only acting in subsequent lessons, until such information is available. By avoiding to base decisions on biased evidence, the teacher would be lessening the chances of producing invalid conclusions about the success of instruction with harmful consequences for students (see Airasian, 2000).

3.4.4.2 Feedback

In the new paradigm, the basic issue with feedback is that it should provide learners with constructive guidance about how to improve (ARG, 2002). Feedback is thus about the promotion of a culture of success where students can build achievements on their previous performance without any comparison with others (Black et al., 2003). This understanding is in line with Ramaprasad's (1983) argument that feedback is actually feedback only if it satisfies the basic condition laid down in his definition:

Feedback is the 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.
(p. 4)

According to this definition, should someone discover that there is a gap but still has no idea about the nature of the discrepancy between actual and desired performance, then that information – which is almost inevitably norm-referenced – fails to qualify as feedback as it does not help him or her to close the gap. Such a process is better described as simply ‘monitoring’ (Wiliam & Black, 1996). On the contrary, for assessment information to count as feedback, it must indicate the existence of a gap between actual and desired levels of performance, as well as suggest actions that prove successful in closing the gap (Wiliam & Black, 1996; also Black et al., 2003). This requirement, which puts feedback at the service of learning, is linked in turn to the concept of student agency as a part of having each student become an independent learner. Building on Ramaprasad’s notion of feedback, Sadler (1989) argues in fact that in order for the student to improve, he or she must: (i) have a notion of the desired standard or goal; (ii) be able to compare the actual performance with the desired performance; and (iii) engage in appropriate action to close the gap between the two. This can only happen, however, if teacher’s standards are available to the student and teacher feedback allows the student to reach these standards (Sadler, 1989). The embedded understanding that feedback should constitute episodes of learning that enable students to connect aspects of poor performance to specific remedial actions encourages me to argue, like Stefani (1998), that what students require is user-friendly information that relates to how they are doing and how specifically they might be able to improve upon this. In this respect, Wiggins (1993; cited in Stefani, 1998) provides teachers with some valid indicators of what it might mean to provide good feedback:

- define the requirements of each learning task;
- describe clearly how performance will be measured/graded/assessed, preferably involving students in this process;
- provide well-articulated descriptors or exemplars of different levels of attainment;
- provide feedback about individual performance expressing this in accordance with agreed criteria;
- relate various aspects of poor performance to specific remedial actions.

These indicators are as much about the communication of expectations (i.e., feedforward) as they are about the communication of progress towards goals (i.e., feedback). In either case, the emphasis is on all students (as opposed to the current practice of having the better students receiving more feedback) receiving

feedback that has high-communication value, in the sense that it can be understood and used by them (Stiggins & Conklin, 1992). This clearly excludes, in spite of many teachers still believing otherwise, that a grade and a short series of comments, usually of a simple praise or blame nature, constitute feedback (Stefani, 1998). On the other hand, if feedback is to provide teachers, as it should, the ‘vehicle for personal dialogue with each learner’ (see Black et al., 2003), it must be given regularly and whilst still relevant, and should also focus on and be task specific (Crooks, 1988). In particular, given that students’ self-perception as learners depends on the quality of feedback they experience over time (Black, 1998), it is essential for feedback to direct attention to the task rather than the learner, as this would lessen the likelihood that the less successful students see it as another confirmation of their inability to perform, yet a further blow to their already low self-esteem (see Wiliam, 1998; also ARG, 2002).

3.4.4.3 Grading

According to Harlen et al. (1992), there are two main ways in which teachers produce summative information about students. These are ‘summing up’ and ‘checking up’:

The former is some form of summary of information obtained through recording formative assessments during a particular period of time and the latter the collection of new information about what the pupil can do at the end of a period of time, usually through giving some form of test. (p. 222)

The possibility that teachers ‘sum up’ formative information for summative purposes – which indicates that these two forms of assessment are not mutually exclusive (see Torrance & Pryor, 1998) – adds credibility, I find, to my position that, within the new paradigm, it is not incompatible for teachers in systems dominated by assessment for selection and certification purposes to have a summative role (see section 3.2.2). For me, instead, the real issue in similar circumstances is to ensure quality in the summative reporting procedures. Given that, for the foreseeable future, teachers will continue, at the most formal level, to judge and to communicate information about student performance through grading (see Airasian, 2000; Brookhart, 1999), my plea for quality summative reporting is basically, at least for the time being, a call for quality grading. My position is that, as long as teachers are required to grade students, we must make sure that grades carry ‘real meaning’ and be appropriate for the

purposes to which the users of their information will put them (see Brookhart, 1999). Airasian (2000) identifies four such purposes: (i) *Administrative* – schools need grades to determine things such as suitability for promotion; (ii) *Informational* – grades are used to tell parents, students, and others about a student’s academic performance; (iii) *Motivational* – the promise of a high grade is used to motivate students to study; and (iv) *Guidance* – grades are used to guide students and parents choose appropriate courses and course levels, and then by schools to sanction or veto these choices.

Although these purposes can be satisfied, almost invariably better, by alternative assessment means, such as profiling and portfolios, that are more conducive to learning than grading, there is little doubt that these new approaches have so far failed to leave a lasting impact on most educational systems (see, for example, Grima & Chetcuti, 2003; Murphy & Torrance, 1988; Weeden et al., 2002). Until such time, there is however an urgent need to improve grading in order to protect the classroom-learning environment. I say this in the knowledge that grades tend to encourage cheating and can negatively influence students’ motivation and self-esteem when lower than expected or consistently low (see Airasian, 2000). Grades may also reward rote learning and foster competitive and grade-hunting attitudes (see Mavrommatis, 1997). Moreover, when grading is cumulative as part of a continuous assessment system, such as when each attempt or piece of work submitted by a student is scored and the scores are added together at the end of the course, students may develop a ‘not for learning’ mindset that it is only worth doing work that contributes to the total (see Sadler, 1989).

Loyd and Loyd’s (1997) four grading principles offer, in my view, a sense of direction towards an enhanced grading process. These are: (i) the grading system should be clear and understandable; (ii) the grading system should be communicated to all stakeholders; (iii) grading should be fair to all students; and (iv) grading should support, enhance, and inform the instructional process. Things are however unlikely to improve unless, contrary to what happens at present, all teachers start getting formal training in grading and are provided with proper guidance about grading policies and expectations (see Airasian, 2000). This development may help, for instance, to change the practice, reported in a number of studies, of having grading being almost solely based on academic evidence of student achievement, with non-academic evidence (e.g., effort and improvement) as a basis for adjustment in student grades, not as the

central determiner of grades (see Airasian, 2000). Although this probably reflects teachers' preoccupation with assigning marks that are publicly defensible (see Peterson & Stack, 1998; also section 3.3.5), it neither does justice to the complexity of the processes involved nor does it lend itself to support learning. On the other hand, useful grading not only draws on several different types of relevant and valid information that gives students more opportunity to show what they can do (Airasian, 2000), but is also accompanied by specific teacher comments about the strengths and weaknesses of a student's work (Mavrommatis, 1997). By adding his or her interpretation to the formal reporting of results, the teacher can put the results in context, identify progress, explain difficulties and indicate ways in which fellow teachers, students, parents and employers can use the information creatively and maximally (Eggleston, 1991).

3.5 Looking Forward with Hindsight

In the last two decades or so, our thinking about educational assessment in general and classroom assessment in particular has come a long way in response to our evolving understanding of what learning entails. We now speak in fact about the emergence of an alternative assessment paradigm that links assessment to helping students learn as opposed to the traditional emphasis on linking assessment to classifying and grading students. From my point of view, though, the problem with these theoretical developments is twofold. First, although they enjoy the support of policy making bodies (see, for example, section 2.6.3), they have failed so far to actually capture the imagination of teachers, students, school administrators, parents and other interested individuals. Second, although we continue to dish out matching, even if laborious, alternative assessment practices, such as records of achievement, profiling and now portfolios, no real effort has been made to create the right environment for their successful implementation. My fear is that to continue ignoring the practitioners and to keep adding to our tried and found lacking list of innovations would risk throwing overboard the whole assessment reform with tragic consequences for learning. As argued in section 3.3.3.3, to dispel this possibility from becoming an uncomfortable reality, we must embark on a well-targeted and marketed educational campaign aimed at promoting the centrality of learning in education and how curricular, pedagogical and assessment practices need to be reorganised to serve this end. This would hopefully be the true beginning of a new educational era.

Part THREE

THE METHODOLOGY

CHAPTER 4

Exploring the Issues

4.0 Dealing with Subjectivity

I described in section 1.1.2 how I underwent a personal paradigm shift in the initial phase of the research and how this seemingly ‘disabling’ internal struggle eventually permitted me to move ahead. This was when I became acutely aware that my subjectivity, which Peshkin (1988a) defines as “an amalgam of persuasions that stem from the circumstances of one’s class, statuses, and values interacting with the particulars of one’s object of investigation” (p. 17), was shaping my inquiry and its outcomes, and that no matter how hard I tried I could never escape it or its consequences. When I subsequently also realised that subjectivity is actually ‘virtuous’, as it is the basis on which a researcher can make a distinctive contribution (Peshkin, 1988a), it became a matter of personal as well as academic integrity to acknowledge my subjectivity to the reader. To be able to do this, I actively sought out my subjectivity by means of a formal, systematic monitoring of my ‘self’. This internal process led me to reflect on the issues of ontology (What kinds of things really exist in the world? What is there? What is the nature of reality?), epistemology (How is it possible, if it is, for us to gain knowledge of the world? What is it to know anything? What is the relationship between the knower and the known?), and methodology (On what rationale is getting ‘knowledge’ based?).

4.1 Reflections on Ontology, Epistemology and Methodology

Endowed with my subjectivity, as a researcher I approach “the world with a set of ideas, a framework ... that specifies a set of questions ... that are then examined ... in specific ways” (Denzin & Lincoln, 1994, p. 11). This research consequently reflects my perspectives on ontology, epistemology and methodology – three dimensions that are intrinsically connected, as “the capacity of whatever methods or procedures give us knowledge of what there is must depend, in part, upon what there is to be known about” (Hughes & Sharrock, 1997, p. 5).

4.1.1 Ontological Reflections

My prior strong belief in ‘objective knowledge’ (see section 1.1.2) has been replaced by an increasing awareness of the ‘multiple realities’ that coexist. My present views are best explained by what Lincoln and Guba (1985) call ‘constructed reality’:

... it is dubious whether there is a reality. If there is, we can never know it. Furthermore, no amount of inquiry can produce convergence on it. There is, in this ontological position, always an infinite number of constructions that might be made and hence there are multiple realities. Any given construction may not be (and almost certainly is not) in a one-to-one relation (or isomorphic with) other constructions of the same (by definition only) entity. The definition is implied by the use of some common referent term, which is nevertheless understood (or constructed) differently by different individuals (or constructors). (pp. 83-84)

I find that Lincoln and Guba’s (1985) courtroom example wonderfully encapsulates the notion of ‘constructed reality’. They point out that during proceedings in a court of law, the defence and the prosecution (using the same events, persons and objects) construct at least two separate realities for the jury’s consideration. This shows how there are multiple ways in which ‘reality’ may be constructed, and how there are multiple rationales for doing so. ‘Truth’ is therefore that which is understood. Thus, even though events, persons and objects are tangible things, the meanings and wholeness derived from or ascribed to such tangible phenomena in order to make sense of them, organise them, or reorganise a belief system, are but constructed realities. Moreover, the sum of these constructed realities will never represent the wholeness of these tangible realities.

This social constructivist vision denies the existence of an all-encompassing god’s eye view of things, one that is above partialities and perspectives. As my world is now bereft of certainties, I have also learned to adopt a cautious approach to all questions of knowledge. I now understand that the natural and the social sciences are partial, as they are both governed by our social and investigative interests (Hughes & Sharrock, 1997). Moreover, I believe that we can never know anything beyond how things appear to us, and that the purpose of research must consequently be to describe the phenomena in our experience and the relationships amongst them, not to speculate about some reality

beyond that experience (Hammersley, 1995). In reality, the researcher can “at best ... feel that he has advanced his problem along an infinite path ... there is no final accumulation and no final solution” (Vidich & Bensman, 1968, p. 396; cited in Peshkin, 1993). Given this progressive, processual nature of research, my quest for knowledge can only be after ‘understanding’, not ‘truth’ or ‘reality’. In order to be fair with the research participants, I must ensure in turn that my portrayal of ‘knowledge’ also respects and represents their personal constructions, irrespective of how informed and sophisticated these may be (Lincoln & Guba, 1990).

4.1.2 Epistemological Reflections

My ontological position leads to a number of important epistemological assumptions. According to Maykut and Morehouse (1994), these include:

1. The knower and the known are interdependent: The way we understand the nature of reality directly affects the way we see ourselves in relation to knowledge. Hence, if we see knowledge as constructed, then the knower can never be totally separated from what is known. Knowledge is seen as being co-constructed.
2. Values mediate and shape what is understood: As the knower and the known are inseparable, the researcher’s values become embedded in the research, in the topic chosen for examination, and in the way the researcher examines the topic.
3. Underlying relationships are complex and multidirectional: As events are mutually shaped, multidirectional relationships can be discovered within situations. The complexity of things is evident from attempts at explanation positions that are interconnected and multidirectional.
4. Only tentative explanations bounded by time and place are possible: Research findings have to value context sensitivity. The intention is to understand a phenomenon in all its complexity and within a particular situation and environment.

5. Research seeks to discover propositions: Research within this approach is characterised by a close examination of people's words, actions and documents in order to discern patterns of meaning that come out of this data. The finality of this is the discovery of propositions that emerge from the patterns of the examined data.

Researchers who embrace these assumptions “stress the socially constructed nature of reality, the intimate relationship between the researcher and what is studied, and the situational constraints that shape the enquiry” (Denzin & Lincoln, 1994, p. 4). The scope of my research – which ultimately seeks to understand how social experience is created and given meaning – is thus to construct a social reality that is based upon the actors' frame of reference within the setting (Lincoln & Guba, 1985).

4.1.3 Methodological Reflections

The experimental research model, so strongly characterised by the non-influential researcher and passive objects of study of the physical sciences, does not make sense when the aim is to get knowledge on/for/with other people who have ‘agency’:

... human beings have agency. Unlike the objects of research in the physical sciences ... human beings are not simply passive subjects of research. All human beings react to situations, including the situations of being researchers or research subjects. In other words, they have agency: they can and do construct interpretations of events, and they can and do use such interpretations as reasons to act in particular ways. (Griffiths, 1998, p. 36)

As “people's behaviour is not caused in a mechanical way, it is not amenable to the sort of causal analysis and manipulation of variables that are characteristic of the quantitative research inspired by positivism” (Hammersley & Atkinson, 1995, p. 8). The consequent futility of trying to discover ‘laws’ of human behaviour directs my study towards interpretive enquiry that, as Hollis (1994) explains, has the capacity to understand the social world from within as opposed to explaining it from without. The underlying belief of the interpretive stance is that “Knowledge of persons [can] only be gained through a ... procedure grounded in the imaginative recreation of the experiences of others to grasp the meaning which things in their world have for them” (Hughes & Sharrock, 1997, p. 98). Understanding social phenomena thus requires that

the lived experiences of others be grasped through the apprehension of the thoughts and understandings that had gone into their production. But, in truth, the researcher can only understand this world of subjective meanings by interpreting it with the knowledge that his or her interpretation is but a construction of the constructions of the actors he or she studies (Schwandt, 1994).

The researcher tries throughout this process to understand the complex world of lived experience from the point of view of those who live it. His or her main goal is an abiding concern for understanding meaning, for grasping the actor's definition of a situation, for *Verstehen* (Schwandt, 1994). The *Verstehen* doctrine, which upholds much qualitative research, dwells in fact on "the human capacity to know and understand others through empathic introspection and reflection based on direct observation of and interaction with people" (Patton, 1990, p. 57). This perspective, which has profound implications on how to study human beings, presumes that the researcher indwells in a situation. This "being at one with the persons under investigation" (Maykut & Morehouse, 1994, p. 25) makes it possible to know the whole, rather than just the pieces:

The pieces of the puzzle are essential to knowing the whole, but in order to gain an understanding of the whole, we must experience, rather than attend to, these pieces, thus allowing the whole to emerge from the experience. (Maykut & Morehouse, 1994, p. 32)

In spite of indwelling, we can still only approximate our understanding of another person's world, as we can never have direct knowledge about it (Maykut & Morehouse, 1994). But although there can be no guarantee that the point of view of the participant will be conveyed in total transparency (Baszanger & Dodier, 1997), the search for 'holistic' knowledge still relies on the empathy that develops from personal contact with people. The promotion of empathy calls in turn for qualitative approaches (such as participant observation, depth interviewing, detailed description, document analysis, and case studies). These give the researcher an empirical basis for describing the perspectives of others whilst also legitimately reporting his or her own feelings, perceptions, experiences, and insights as part of the data (Patton, 1990). This methodological approach ultimately requires a strong measure of reflection on the

researcher's part. The researcher, who is part of the investigation as participant observer, needs in fact to 'distance' him or herself from the situation to rethink the meanings of the experience.

4.2 Selecting the Research Design and Methods

My declared position in favour of qualitative research approaches should not be taken to mean a total negation of quantitative modes of enquiry, as these may also be used to elucidate a research process that is essentially qualitative (see Hanafin, 1995). Consequently, although I fully appreciate the distinctions and differences between the respective philosophical bases of the qualitative and quantitative research paradigms, I am committed to "use whatever information is available" (Hammersley, 1992, p. 199).

4.2.1 Opting for a Qualitative Case Study

The research design depends on what the researcher wants to know (Merriam, 1998). In my case, the search was for a research design that respects my value positions (see section 4.1), is an appropriate means for illuminating the phenomenon being investigated (i.e., PMI teachers' classroom assessment practices), and takes due consideration of the impact of the research setting (i.e., the Junior College). My research study is particularly suitable for case study as it seeks to understand humans engaged in action and interaction within a particular context (Collins & Noblit, 1978; cited in Merriam, 1998), is concerned with processes (Merriam, 1998), and asks 'how' and 'why' questions (Yin, 1994). Although 'case study' is a generic name that has a range of meanings to different authors (see Bassey, 1999; Merriam, 1998), my understanding of the term is largely guided by Merriam's (1998) definition:

[I] see the case as a thing, a single entity, a unit around which there are boundaries. I can 'fence in' what I am going to study. The case then, could be a person such as a student, a teacher, a principal; a programme; a group such as a class, a school, a community; a specific policy; and so on. (p. 27)

This encapsulates the awareness that "the single most defining characteristic of case study research lies in delimiting the object of study, the case" (Merriam, 1998, p. 27)

that is in turn “a phenomenon of some sort occurring within a bounded context” (Miles & Huberman, 1994, p. 25). Although it is epistemologically impossible to give an exhaustive account of any object (Hammersley & Atkinson, 1995), a case study methodology still provides “a means of investigating complex social units consisting of multiple variables of potential importance in understanding the phenomenon. Anchored in real-life situations, the case study results in a rich and holistic account of a phenomenon” (Merriam, 1998, p. 41).

A case study design does not claim any particular method for data collection or data viewing (see Stake, 1994). But the underlying complexities of the social world of teachers – which lies at the heart of this study – are better explored by a qualitative case study that is able to connect ordinary practice in natural habitats to the abstractions and concerns of diverse academic disciplines (Stake, 1994). Even so, the choice amongst the various types of qualitative case study designs remains vast. The very nature of my research problem eventually directed me towards a case study that is ‘descriptive’ with reference to ‘special features’, ‘interpretive’ in its ‘overall intent’, and leaning towards the ‘ethnographic’ as far as ‘disciplinary orientation’ goes (see Merriam, 1998). It is ‘descriptive’ in the sense that my write-up – which describes, elicits images and analyses situations – offers a rich portrayal of the phenomenon under study that attests to the complexities of the situation. The resulting rich, thick descriptions are the backbone of my ‘interpretive’ case study. My willingness to gather as much information about the problem as possible placed me in a position to analyse, interpret and theorise about the phenomenon. In the process, I have been able to get close to my participants within their natural setting. This is where the semblance of ‘ethnography’ comes in. For, in the field, I have been able to observe them directly, was in a position to gather relevant documents, and had access to their subjective factors (i.e., thoughts, feelings, and desires) through deep interviewing.

4.2.2 Using Multiple-Methods within a Flexible Research Framework

My decision to spread the ‘net of evidence’ widely resulted from the awareness that “every method of data collection is only an approximation of knowledge. Each provides a different and usually valid glimpse of reality, and all are limited when used

alone” (Warwick, 1973, p. 190; cited in Peshkin, 1993). In order to gain a comprehensive understanding of the research phenomenon, I thus chose to use the complete set of qualitative data collection methods mentioned by Patton (1990). These include: (i) in-depth, open-ended interviews; (ii) direct observation; and (iii) written documents. In all three methods, the investigator is the primary instrument for gathering data and has to rely on his or her skills and intuition. Human instrumentation offers the distinctive advantage that as the researcher responds to the context, he or she can adopt techniques that fit the circumstances, thus managing to keep the total context in consideration (Guba & Lincoln, 1981; cited in Merriam, 1998). My chosen methods also share important methodological characteristics. Whether the researcher is interviewing, observing or analysing documents, he or she tracks down leads, remains open to new insights, and is sensitive to the data (Merriam, 1998). This helps the study to remain ‘open’, as in the ethnographic tradition, to elements that cannot be codified at the time of the study (Baszanger & Dodier, 1997).

My field presence has been particularly inspired by ethnographic researchers who typically participate “in people’s lives for an extended period of time, watching what happens, listening to what is said, asking questions – in fact, collecting whatever data are available to throw light on the issues that are the focus of the research” (Hammersley & Atkinson, 1995, p. 1). In this active approach to research, which signals a shift away from the more traditional ‘armchair theorising’ approach, the researcher has to suspend a wide range of common sense and theoretical knowledge in order to minimise the danger of taking on trust misleading preconceptions about the setting and the people in it (Hammersley & Atkinson, 1995). It is important here to distinguish between ‘preconceptions’, which should be avoided, and what Malinowski (1984) calls ‘foreshadowed problems’, which characterise quality research:

Good training in theory, and acquaintance with its latest results, is not identical with being burdened with ‘preconceived ideas’. If a man sets out on an expedition, determined to prove certain hypotheses, if he is incapable of changing his views constantly and casting them off ungrudgingly under the pressure of evidence, needless to say his work will be worthless. But the more problems he brings with him into the field, the more he is in the habit of moulding theories according to facts, and of seeing facts in their bearing upon theory, the better is he equipped for the work. Preconceived ideas are pernicious for any scientific work, but *foreshadowed problems* are the main endowment of a

scientific thinker, and these problems are first revealed to the observer by his theoretical studies. (Malinowski, 1984, p. 9; cited in Stake, 1994)

The case content itself actually evolves as the study progresses. In reality, “one cannot know at the outset what the issues, the perceptions, the theory will be. Case study researchers enter the scene expecting, even knowing, that certain events, problems, relationships will be important, yet discover that some are of little consequence” (Stake, 1994, p. 240). This reveals the impracticality of working within a static research design burdened by rigid methodological rules. On the other hand, by opting for a primarily inductive research process, I have been able to build towards a theory from observations and intuitive understandings gained in the field (Merriam, 1998). In the knowledge that my manner of ‘being’ within the setting not only influences the nature of the investigation, but also the meanings I confer to the data gathered, I also decided to record, reflect upon, and integrate within the narrative my ‘being at the setting’.

4.3 Presenting the Data Collection Methods

I lacked documentary evidence where it mattered most – that is, at classroom level. This made it even more vital that I consider carefully what role to delineate to observations and interviews, the two “mutually reinforcing qualitative techniques” (Patton, 1990, p. 32). As data from each can be used to illuminate the other (Hammersley & Atkinson, 1995), I would have liked to give prominence to both in my study. However, as I will shortly explain, I eventually shifted the emphasis on interviews for mainly practical reasons. Thus, although I have used all the data collection methods normally associated with qualitative research (i.e., interviews, observations and document analysis), my main data source has been interviewing. This testifies that within specific studies, more than grand methodological theories, it is the prevailing circumstances that suggest the more effective methods to be employed.

4.3.1 Documents

My study takes place in what Atkinson and Coffey (1997) call a ‘literate society’ – the setting is a school in and about which documents are written, read, stored and

circulated. Although such documents cannot be treated as accurate portrayals of reality (Hammersley & Atkinson, 1995), it is by giving attention to them that a case study can be grounded in the context of the problem being investigated (Merriam, 1998). There is however some disagreement about their eventual role during data analysis and interpretation (see Atkinson & Coffey, 1997). Indeed, whilst some researchers see documents as data in their own right, others simply use them to support or validate other data, such as interviews and observations. In my view, the role of documents during this decisive phase should depend primarily on the trustworthiness of the researcher, on how systematic, rigorous and comprehensive their collection has been. Thus, my yearning to produce an account that does justice to the setting it purports to describe (Atkinson & Coffey, 1997) led me to collect all kinds of school-related documents. The written sources I consulted may be grouped under published and non-published materials. The published documents were mainly newspaper articles, University of Malta press releases, MATSEC syllabi and information materials, and Education Division and Junior College publications. Being a school insider, I also had easy access to unpublished internal memos, reports and available records.

The difficulty I encountered here was that whilst I found myself immersed in documents that provide rich historical and contextual information, I only met with scarce documentation at departmental and teacher levels. This difficulty came particularly to the fore when my research focus began to shift away from the evaluation of the PMI course to dwell more closely on the personal experiences of teachers (see section 1.2.1). This lack of documentary evidence is however data in its own right that tells us something about the research context (Guba & Lincoln, 1981; cited in Merriam, 1998). With regards to documentation at classroom level, I began by collecting the teachers' available documents such as class records, worksheets and class tests. I subsequently supplemented these by what Merriam (1998) calls 'researcher-generated documents'. These are documents that are produced either by the researcher or for the researcher by participants after a study has already begun. Examples of the former in my study are the classroom layouts that I recorded during classroom observations. In the latter case, I asked teachers to predict students' performance in examinations. The idea behind researcher-generated documents is to include in a study alternative material traces that benefit "qualitative researchers who

wish to explore multiple and conflicting voices, differing and interacting interpretations” (Hodder, 1994, p. 395).

4.3.2 Observations

When observation involves participation in the social process, it is possible for the observer to accumulate not only ‘explicit knowledge’ and ‘focal awareness’ of events that any observer could obtain, but also the ‘tacit knowledge’ and ‘subsidiary awareness’ which could only come from direct experience (Pollard, 1985).⁵ But for the participant observer to be sufficiently part of the situation to be able to understand personally what is happening (see Patton, 1990), he or she needs to have considerable ‘research time’ at his or her disposal – something that I lacked during the larger part of the study. The only period throughout my three years in the field when I was actually at school and not teaching full-time was from January to March of Year 3 – a three month period during which I was very busy conducting limited classroom observations and the second and final phase of formal interviewing (see section 5.2.5). As I conducted the study at my own school, participation was ‘real’ for me, not just a research strategy.

I realised early in the study that my heavy teaching schedule and administrative work would limit my ‘informal’ field interactions mainly with those participants who either share my office or who are in my social group (see section 5.2.2). This meant that with a good number of possible participants I would only be able to interact rarely and for a restricted period of time. For whilst I only have chance and fleeting encounters with these colleagues, significant interaction with them was rendered more difficult by the ‘free movement’ policy practised at our school. Teachers can leave school whenever they are not strictly attending to their teaching and related duties. This meant that even

⁵ This terminology derives from Polyani’s work that Maykut and Morehouse (1994) summarise as follows: “Two types of knowledge play a part in the way we understand the world, tacit and explicit knowledge, but tacit knowledge is more basic – it comes before explicit knowledge. Tacit knowledge is unarticulated knowledge; it is unformulated, such as the type of knowledge we have in the act of doing something. Explicit knowledge is that which is or can be written down in words, maps or mathematical formulas. ... Tacit knowledge is gained by indwelling. When one lives within a situation one learns to pay attention to the subsidiary, that is, one learns to attend away from the object and toward the meaning of the object” (p. 31).

if I had the time to establish fruitful contacts with teachers in neighbouring offices, which anyway I did not have, it would still have been very unlikely that I would find anyone there who was not busily attending to some administrative duty or other. In view of these interactive restrictions at school, any attempts on my part to observe most of the participants outside the classroom, although clearly purposeful, could not hope to be systematic. As I could not abide by the observational canon of studying my surroundings regularly and repeatedly (see Adler & Adler, 1994), I chose to rely mainly on formal interviews for my data. But in the field I still considered myself and behaved, within the limits imposed by my 'real' teaching duties, as a participant observer who

... is fully engaged in experiencing the setting under study while at the same time trying to understand that setting through personal experience, observations, and talking with other participants about what is happening. ... In participant observation the researcher shares as intimately as possible in the life and activities of the setting under study. (Patton, 1990, p. 207)

In reality, there are different 'social roles' that participant observers may adopt in the field. They may even in the course of the same study change from one role to another (Burgess, 1982b) – something that I did. Indeed, with the exception of classroom observations, where I was a passive onlooker, all other observations were conducted whilst I was participating within the department in my role of teacher. My initial intention was to adopt a participant observation style that involves "a formal participatory role in a social situation without the emotional involvement that normally accompanies participation" (Gans, 1982, p. 54). I feared that should I fail in this respect, I would risk the danger of over-rapport that leads to identifying with participants' perspectives, and hence the failure to treat them as problematic (Hammersley & Atkinson, 1995). In the field, I sought to address this fear by adopting what Gold (1958; cited in Burgess, 1982b, Burgess, 1984, Hammersley & Atkinson, 1995, and Merriam, 1998) calls the 'complete participant' role. This involved banking on my natural immersion in the setting as a teacher to conduct covert observations. However, in recognition of the inalienable influence of my subjectivity and following ethical and methodological rethinking on my part (see section 5.1.2), I later adopted Gold's more congenial 'participant-as-observer' role. Thus, once I formally negotiated access, the fully informed and consenting participants not only knew about my

observations, but they also realised that, apart from collegial and friendship interests, my interactions with them were also guided by my research interests. Burgess (1982b) warns that in order not to over-identify with the researched, it is essential for the participant-as-observer to remain a ‘stranger’ whilst involved in the situation under study. In my case, more than ‘remain’, I had to endeavour to ‘become’ a stranger in my own natural setting. Thus, whilst I continued to participate in the life of my department, I tried to limit my involvement so that I could still function as a researcher. In order to do this, I had to actively cultivate a degree of skill in retaining what Pollard (1985) calls ‘detachment in my head’.

In spite of my efforts to record action as it unfolds and to discuss it with participants, I admit to not having been repeatedly that close to a number of participants in a range of situations that would satisfy the academic canons of observation data. Notwithstanding this, my observational data have still benefited my study in at least two important ways. First, they allowed the research problem to emerge from the field (see section 1.2.1). Understandably, as my research focus began to sharpen, so did my observational focus. My most selective observations – namely, those inside the classrooms – bring me to the second benefit. My classroom observations, above all other observational data, served to provide the context and specific incidents that I used as reference points during the subsequent interviews. Having first observed them in class for a week, I was able to discuss with PMI teachers their views with regards to specific behaviours and episodes that I had witnessed. This ‘positioning’ helped in no small measure to contextualise, and thus deepen, my understanding of teachers’ classroom assessment practices.

4.3.3 Interviews

Apart from the limitations I encountered with documents and observations, my decision to favour interviewing as the primary mode of data collection also depended on its adequacy for the purposes of this research. For when the researcher is interested (like I am) in the participants’ experience and what meaning they make out of that experience, interviewing may then be, in most cases, the best avenue of inquiry (Seidman, 1998). Interviewing – which normally entails a “person-to-person encounter

in which one person elicits information from another” (Merriam, 1998, p. 71) – assumes that “the perspective of others is meaningful, knowable, and able to be made explicit” (Patton, 1990, p. 278). Apart from providing access to the participants’ present feelings, thoughts and intentions, interviews can also bring the researcher in contact with past events or events that are out of his or her reach (Patton, 1990). Gaining understanding of human beings through interviews can be achieved in a variety of ways that differ along a continuum with highly structured, questionnaire-driven interviews at one end and unstructured, open-ended, conversational formats at the other (Merriam, 1998). The structured interviewing at one end refers to a situation in which the interviewer asks each respondent a series of pre-established questions with a limited set of response categories, and there is generally little room for variation in response except where an infrequent open-ended question may be used (Fontana & Frey, 1994). At the other end, where the interview situation is far from defined, what Burgess (1982a) calls ‘non-directive interviews’ even allow respondents to take the subject of discussion in whatever direction they prefer.

4.3.3.1 Deciding on the Interview Format

Two important considerations guided my choice of interviewing style. First, the interview had to reflect the mode of understanding implied by qualitative research. More precisely, it had to reflect my constructivist persuasions of the interviewer and interviewee actively constructing some version of the world that is appropriate to the person being interviewed and the context in question (Silverman, 1993). The understanding that knowledge is constructed through the interaction of interviewer and interviewee (Kvale, 1996) bars researchers from continuing to interview persons as ‘objects’ with little regard for them as individuals. This traditional hierarchical situation in interviewing thus needs to be replaced by a closer relation between the interviewer and interviewee, a relation that attempts to minimise status differences (Fontana & Frey, 1994). As my primary aim was to generate data that give an authentic insight into people’s experiences (Silverman, 1993), I could not go for an interview format that seeks to capture ‘precise’ data of a codable nature. An interview that seeks to explain behaviour within pre-established categories would have actually made it more likely that I get reactions from participants to my preconceived notions of the

world, rather than access their perspectives and understandings of it (Merriam, 1998). My search for meaningful relations to be interpreted is better served instead by an interview format that does not impose *a priori* categorisations that may delimit the field of inquiry (Fontana & Frey, 1994). Basing myself on these reflections, I came to the conclusion that all structured interviews (not just the questionnaire-driven ones that are inherently quantitative in nature) are not compatible with qualitative research.

My second consideration was my inexperience with qualitative interviewing. As a beginner, I felt threatened by the openness and uncertainty of the less structured interview (Maykut & Morehouse, 1994). I feared that without the necessary structures, I would have problems concerning data quality, control and even time management. These fears – combined with my desire to still benefit from what the less structured interview has to offer and to somewhat balance, not invert, the asymmetrical power relations found in most interviewing situations – gradually channelled me towards the middle ground position of the semistructured interview:

In this type of interview either all of the questions are more flexibly worded, or the interview is a mix of more and less structured questions. Usually, specific information is desired from all respondents, in which case there is a highly structured section in the interview. But the largest part of the interview is guided by a list of questions or issues to be explored, and neither the exact wording nor the order of the questions is determined ahead of time. This format allows the researcher to respond to the situation at hand, to the emerging worldview of the respondent, and to new ideas on the topic. (Merriam, 1998, p. 74)

I saw in semistructured interviews, which are typically referred to as ‘depth’ or ‘in-depth’ interviews (Lincoln & Guba, 1985), an interviewing format that is both flexible and controlled. This would give me an opportunity “to probe deeply, to uncover new clues, to open up new dimensions to the problem and to secure vivid, accurate, inclusive accounts” (Burgess, 1982a, p. 107). Although such an interview appears to proceed like a normal conversation, the researcher still defines and controls the situation (Kvale, 1996) by developing and sequencing questions with the aid of an interview guide. This guide, which contains a list of questions or issues to be explored in the course of an interview, is prepared so that basically the same information is obtained from a number of people by covering the same material (Patton, 1990).

Although I readily concede that semistructured interviews constitute what Fontana and Frey (1994) term a ‘one-way pseudo-conversation’, I do not concur with their assertion that this leads to a one-sided and therefore inaccurate picture. I hold on the contrary that when the interviewer indulges in personal opinions – which would include answering questions raised by interviewees and providing personal opinions – he or she would be undermining his or her neutrality concerning what the interviewee is saying. This failure to adopt a neutral stance may actually dent the interviewer’s rapport with interviewees, as this is built on the ability to convey empathy and understanding without judgement (Patton, 1990). Sincerity in response is instead perhaps best guaranteed when the interviewee feels comfortable and confident in talking with the interviewer – something that develops over a period of time (Whyte, 1982). The prolonged engagement (sometimes over a number of sessions) that characterises in-depth qualitative interviews in fact benefits understanding as it allows the interviewer to establish rapport and a climate of trust with the interviewees (Maykut & Morehouse, 1994). A good beginning in this direction is for the interviewer to be respectful, non-judgemental, and non-threatening (Merriam, 1998).

4.3.3.2 Qualitative Interviewing

Interviews need not necessarily be formal affairs – that is, arranged meetings in bounded settings and out of earshot of other people. In fact, interviews in qualitative studies often take place whilst the researcher is a participant observer (Maykut & Morehouse, 1994). Irrespective of this, a good qualitative interviewer is someone who is primarily motivated by a deep and genuine curiosity to understand another’s experience in a tactful manner (Maykut & Morehouse, 1994). Characteristically,

... the research interviewer listens more than he talks, and listens with a sympathetic and lively interest. He finds it helpful occasionally to rephrase and reflect back to the informant what he seems to be expressing and to summarise the remarks as a check on understanding. (Whyte, 1982, p. 111)

The interviewer also needs to be a skilled observer, one who is “able to read nonverbal messages, sensitive to how the interview setting can affect what is said, and carefully attuned to the nuances of the interviewer-interviewee interaction and relationship”

(Patton, 1990, p. 32). The value of the interview ultimately depends on the interviewer knowing enough about the topic to ask meaningful questions in a language easily understood by the interviewee (Merriam, 1998). Good interview questions – that is, those that draw the participant into conversation and yield useful information – are typically open-ended ones that are not easily answered with a discrete response, such as ‘yes’ or ‘no’, or a brief word or phrase (Maykut & Morehouse, 1994). The interviewer must reflect at all times upon what is being said, ask him or herself what each statement means and how he or she can best encourage the informant to clarify a certain point or give detail on an item hinted only (Whyte, 1982). This calls for a flexible approach that permits the researcher to follow promising leads or return to earlier points that seem to require fuller development (Lincoln & Guba, 1985).

4.4 Ethical Considerations

What follows is an outline of the ethical philosophy that guided my study. I do not claim that all my field decisions, particularly those in the beginning stages, reflected exactly this view. There have indeed been times when I took decisions without being knowledgeable enough to see their full consequences. To my credit, however, whenever I realised subsequent to a decision that I could have acted in a more ethical manner, I did my best to make up for it. I actually embarked on this study thinking that ‘wanting to gain knowledge without desiring to harm anyone’ was a good enough parameter within which to work ethically. Only later, once I realised the gross naivety of this position, did I begin referring to the literature (including professional codes of ethics) in an effort to formulate my own ethical vision. But rather than relying blindly on some established code of ethics – which does not in any case tell you what to do in specific cases – I sought to come up with my own interpretations (Sowder, 1998). This basing of ethical practice on my own values and ethics (Merriam, 1998) still did not stop me from experiencing difficulties in dealing with ethical dilemmas. In practice, these ethical concerns revolved around issues of harm, consent, deception, privacy and confidentiality of data (Punch, 1994).

As the study progressed, I became ever more conscious that the research relationship between researcher and researched is invariably unbalanced. In particular, the

awareness that the information I obtained from the participants gave me possibly harmful power over them (Simons, 1982) made it imperative that I protect them (Ely et al., 1991). Consequently, whilst I understand that the aim of research is to produce true accounts of social phenomena, I do not agree that this should be pursued at all costs (Hammersley & Atkinson, 1995). As the 'right to know' can easily clash with the principle of respect (Pring, 1984), it is essential for each participant to give his or her informed consent. This requires that he or she is fully informed, is competent to give consent, fully comprehends the conditions of consent, and gives consent voluntarily (Sowder, 1998). The given information should thus include aspects such as the research goals, the participant's role, why and how participants were selected, the risks and benefits of participation, the ways in which the data will be used, and the possibility of withdrawing participation at any time (Sowder, 1998). I understand however that it is not always possible, or advisable, to tell 'all' about the study to the participants:

... at the initial point of negotiating access, the ... [researcher] her- or himself often does not know the course the work will take, certainly not in any detail. But even then, once the research problem and strategy have been clarified, there are reasons why only limited information may be provided to participants. For one thing, ... divulging some sorts of information might affect people's behaviour in ways that will invalidate the research. (Hammersley & Atkinson, 1995, p. 265)

As long as the interests of the participants are protected, I do not see this 'holding back on some information' as deceitful. This permits me to identify with Ely et al.'s (1991) views on access negotiations:

The person making entrée need not feel she or he is being sly or dishonest by making that introduction as general as serves its aims. Some of its purposes are to communicate about the study, the roles of the researcher that are in line with that, the ways participants will be involved and their rights, and the support that is needed, while not providing information that would impinge on the very phenomenon to be studied. This aim is not dishonest *per se*. (pp. 38-39)

Apart from deception, which I relate primarily to being untruthful or misleading, the participant is also not respected when things said and done for private consumption are made public. Ethical research has to strive instead for an appropriate balance between the public 'right to know' and the individual 'right to privacy' (Simons, 1989). The

researcher can achieve this balance by negotiating clearance of information offered by participants and used as data for the study before distribution to anyone else (Simons, 1982). This is where the principle of confidentiality – which practically assigns ownership rights to participants in order to protect private and possibly damaging data from unnegotiated dissemination (Hammersley & Atkinson, 1995; Simons, 1989) – comes in. Confidentiality, which has its most obvious application in relation to interview data, can in principle at least be extended to observational data as well (Hammersley & Atkinson, 1995). As a matter of fact, I offered two types of confidentiality in my study. First, I treated interview data as ‘off the record’ until cleared. And second, I treated all other sources of data as ‘on the record’ unless specifically asked to treat them as confidential. In the knowledge that confidentiality has a much less formidable claim as soon as one moves away from the release of basic data to interpretation and to drawing conclusions (Pring, 1984), I resolved not to offer the participants any form of ownership during these two phases.

The concept of anonymity, which offers individuals some privacy in the research process or protection from identification whilst allowing more explicit discussion or reporting of contentious issues, is often linked with confidentiality (Simons, 1989). But in spite of the strong feeling amongst fieldworkers that settings and respondents should not be identifiable in print (Punch, 1994), anonymisation appears almost impossible to achieve, especially in an internal study (Simons, 1989). Punch (1994) puts forward some possible reasons for this:

Pseudonyms can often be punctured by looking up the researcher’s institutional affiliation at the time of the project. ... In addition, the cloak of anonymity for characters may not work for insiders who can easily locate the individuals concerned or, what is even worse, *claim* that they can recognize them when they are, in fact, wrong. [Again] many institutions and public figures are almost impossible to disguise. (p. 92) (emphasis in original)

Nevertheless, it may still be important to maintain the principle of anonymity when reports are disseminated, as this decreases the likelihood of identification over time and distance (Simons, 1989). This also applies to unpublished PhD studies, as these are usually public documents lodged in a library and open to all. To protect further the participants from harm, the circulation of the report can be limited to ‘relevant

audiences' that are negotiated (with the proviso that these may need to be renegotiated during the study itself) and declared from the outset (Simons, 1989).

Exploitation – which occurs when participants get little or nothing in return for supplying the researcher with information (Hammersley & Atkinson, 1995) – is closely linked to the issue of harm. The widespread practice of having research serving only the personal advancement of the researcher (Seidman, 1998) is no longer ethically acceptable (Ely et al., 1991). It violates one of the main objectives of the principle of beneficence – that is, the maximisation of benefits to participants. Understanding the need to give something back, my resolve has been to act upon the 'pay back' strategy championed by Ely et al. (1991). This involves telling the participants' stories rather than impose my own, reporting their meanings, and describing their social context, not as separate, but as it is lived and understood by them. Moreover, I can help participants by sharing with them one of the best benefits I have to offer – the knowledge that I acquire (Cassell, 1982; cited in Sowder, 1998).

Finally, it must however be said that, given the complexities involved, it is not easy to conduct educational research that fully respects all the rights of the individuals involved. Even seasoned researchers continually experience difficulties in dealing with ethical dilemmas. Believing that the integrity of the research and the researcher ultimately depends upon the fulfilment of expectations generated by interpersonal perceptions, I have sought to position 'trust' – which I tried to win rather than assume – at the centre of my work (Simons, 1989). Even though no set of rules can fully capture the spirit of trust, I have worked for it by abiding scrupulously to all the obligations that my agreements with the participants, which were meant to cover some of the dangers mentioned in exchange for conceding the right to know, placed upon me (Pring, 1984). Indeed, if I did err in my interactions with the participants, I did so on the side of caution and respect. For I desire my friends and colleagues to judge me well first and foremost as a person, and only then as a researcher.

CHAPTER 5

Implementing the Study

5.0 The Research Site

This research was inspired by my ‘habitual’ presence at the Junior College. It was my ‘being there’ that initially triggered within me a desire to explore the PMI option and eventually catapulted me into the realm of teachers’ classroom assessment practices (see Chapter 1). Doing research as a ‘real’ insider – that is, someone who was internal to the setting prior to and irrespective of the study – thus had little to do, in my case, with being already ensconced in a setting and deciding ‘opportunistically’ to study it (see Reimer, 1977; cited in Adler & Adler, 1994). Still, my ‘continued presence there’ enabled me to bring my complete range of senses to the investigation, to have sufficient time in which to be attentive, and to bring ‘fullness’ to the investigation (Peshkin, 1988b) in spite of having a full-time teaching load throughout most of the study. Apart from these methodological considerations, the Junior College also offered distinctive advantages over all other ‘would have been possible’ sites. It offered me that case from which I could learn most about the issues of central importance to the purpose of the research. Particularly decisive here was my realisation that the Junior College presents what Patton (1990) terms a ‘typical case’ and a ‘critical case’.

The Junior College is a typical site because “it is not in any major way atypical, extreme, deviant, or intensely unusual” (Patton, 1990, p. 173) in comparison to other institutions in which the phenomenon that interests me also occurs. The Junior College, similar to other sixth form colleges in Malta, aims mainly to prepare students for university education, and its study programmes follow closely the MATSEC Matriculation Certificate. Again, all eligible students who apply are offered a placing. This entry policy results in a student population that comes from all over the island, has a good prior school-type background and social mix, and represents a wide span of ability ranges and levels of motivation. Likewise, the teachers at the Junior College offer a good mix with regards to age, gender, qualifications and teaching experience. At the same time, the Junior College is a critical site because it is the “site that would

yield the most information and have the greatest impact on the development of knowledge” (Patton, 1990, p. 174). This is mainly due to its ‘overwhelming numerical presence’ within the local pre-university sector (see section 2.4.2). As a matter of fact, more teachers teach and more students study the PMI option at the Junior College than in all the other local sixth form colleges put together. Moreover, as my research interests lie mainly in exploring and contextualising PMI teachers’ classroom assessment practices, the possibility offered by the Junior College to study a well-assorted group of teachers made this location especially appealing. In other colleges, I would have found only one or two PMI teachers at most.

5.1 The Pre-Fieldwork Phase

5.1.1 Entering the Field

My field experience can be classified under two distinct, consecutive stages – the ‘pre-fieldwork’ phase that was followed by the ‘fieldwork’ phase proper. My pre-fieldwork phase, during which I actually entered the field, somewhat parallels what Corsaro (1980; cited in Lincoln & Guba, 1985) calls ‘prior ethnography’. Here, the researcher becomes a participant observer in a situation for a lengthy period of time before conducting the study in order to sensitise the human instrumentation (Lincoln & Guba, 1985). In my study, this ‘acclimatisation’ phase was however a natural consequence of my insider status, rather than a conscious research strategy on my part. This was when I sought to identify as an insider the ‘bumps’ (see Wolcott, 1994) around me – that is, those elements within the mathematics department that stand out from an otherwise flat landscape. This exercise eventually focussed my attention on the PMI option. Having realised by then that I could not gain ‘true’ understanding simply on the merit of ‘being there’, I began subjecting my school experiences to analytic reflection (see Hammersley & Atkinson, 1995) and embarked on two other research strategies.

My first strategy was to delve into the backgrounds, underlying philosophies and objectives of the Junior College and the Matriculation Certificate. As they were then both educational novelties, my main sources of information were legal notices, commissioned reports, articles and letters in local newspapers, newsletters, and

internal and public communications. Thanks to this ‘unearthing’ process, not only did I enter the field already very well acclimatised to the setting in view of my insider status, but I was also very much aware of the underlying contextual tensions. This awareness proved extremely useful, both in the field as I progressively focused my investigation and during the analysis of data, as it helped me to understand better my participants’ multifaceted perspectives. My second strategy drew on Hammersley and Atkinson’s (1995) consideration that there are no fixed rules for deciding how far the initial research problem has to be elaborated before the collection of data begins. Following my decision to introduce ‘observations’ to supplement my ‘being there’, I began recording ‘unplanned’ observations and pertinent analytical comments in what I, for lack of a better name, initially called ‘observational notes’. I laid down the criteria for the inclusion of records in my first entry:

I intend to record what I see and hear at school that may shed light on the PMI option within the Junior College. I also intend to include my reflections on these observations and related readings. (September, Year 1)

Typical of early data collection stages, these initial observations were unfocussed (Morse, 1994) and dredged up much information that, while interesting, was not exactly relevant (Lincoln & Guba, 1985). A glance at these notes reveals their breadth – I was recording and reflecting upon conversations with colleagues (who did not necessarily teach PMI), departmental meetings, interactions with students, and the MATSEC PMI syllabus panel meetings to which I attended as an observer. Even at this early stage of the research, I was careful to distinguish my analytical comments (i.e., what went ‘inside me’) from my descriptions and the accounts I chanced upon (i.e., what I saw and heard ‘out there’). Later, during the fieldwork phase, I developed these observational notes into a proper fieldwork journal (see section 5.3.1).

5.1.2 Ethical and Pragmatic Rethinking

My pre-fieldwork phase was covert. Throughout this period, I was able to record unsanctioned observations at will mainly by virtue of my natural presence within the mathematics department. But subsequent ethical rethinking on my part (see section 4.4) made me realise that such deceptive and covert practices are not in keeping with

the ethical practices of the alternative paradigm (Maykut & Morehouse, 1994). Even though I had never meant to hurt anyone, I decided at one point that in order to somewhat assuage my ethical insensitivity I would not make any third-party references (say, direct quotes or particular episodes) in my study that predated my formal access to the Junior College as a researcher. This did not affect the quality of my data, as I was able later to come across again all the important evidence I had 'lost on ethical grounds' by observing persistently enough (see Ely et al., 1991). My decision to abandon covert research however also had an equally important pragmatic component. This emerges clearly from the following fieldwork journal entry:

I find my covert status no longer tenable on ethical grounds. I cannot continue with this deception any longer. ... I also need to start collecting data that is more related to 'classroom assessment'. I can no longer rely on 'chancing upon it'. This necessitates that I start approaching and selecting colleagues who are willing to take part in the study. (October, Year 2)

As my covert observations were proving rather limiting, I also desired to go overt to manoeuvre myself into a position from which I could collect the necessary data (Hammersley & Atkinson, 1995). Hoping at the same time to open my study to some of the important multiple roles that participants could play (see Maykut & Morehouse, 1994), I henceforth tried to work 'with' participants rather than 'on' subjects.

5.2 Formally Entering the Field

By the end of the pre-fieldwork stage, my research interests, whilst retaining the PMI option and the Junior College as points of reference, had swung decisively towards classroom assessment. In particular, I found myself becoming increasingly intrigued by the relationships between assessment procedures, learning goals, teaching activities, and learning processes (Cumming & Maxwell, 1999). Thinking vaguely along these lines, I began making preparations to enter formally into the field.

5.2.1 Gaining 'Formal' Access

Initially, I identified two gatekeepers – the Junior College principal and the mathematics subject coordinator. But later, acting upon a suggestion by the principal, I

added the Junior College registrar. My first tangible step towards obtaining official access was to approach ‘formally’ my subject coordinator. The fieldwork journal provides a summary of my requests during a short private meeting with him:

Today I informed the subject coordinator that I am proposing to explore PMI teachers’ classroom assessment practices for my PhD. I explained that it was my intention to keep written records of the things I observe within our department, to interview volunteering PMI teachers, and to make use of school and departmental records. I also mentioned the possibility that I may need to ‘enter’ classrooms at some point or another. Before leaving I asked him to treat our conversation, at least for the time being, as confidential. I explained that once I have official access, I would be contacting each PMI teacher personally. (November, Year 2)

Agreeing to keep the matter confidential for the time being, the subject coordinator suggested that I write an explanatory letter to the principal. He assured me that once I have the approval from the principal’s office, he would help me in any way he could. The next thing I did was to ask for a meeting with the principal and to formulate the access letter suggested by the subject coordinator. When I met the principal in his office, I handed him the letter (which he carefully read) and briefly explained the aims of my research and how I was proposing to achieve them. A couple of days after this brief, cordial meeting, I received an annotated copy of my letter by internal post. At the bottom of my letter, the principal wrote and signed: *‘Approved and encouraged. I am forwarding a copy to the college registrar. You may be needing his help as well’*. The thought of approaching the registrar – who effectively ‘polices’ college records – had not crossed my mind. I contacted the registrar the next day. After I briefed him about my research, we discussed ways in which he and his staff might be helpful. We agreed that, as the need arises, I would have access to the school records as long as I made my demands in writings and as long as these did not violate any confidential information.

Contrary to what happens in a number of studies (see Hammersley & Atkinson, 1995; Lincoln & Guba, 1985), none of my gatekeepers tried, even minimally, to redirect the nature of my study, and the only ‘bargain’ I struck was with the registrar for pure ethical reasons. Given my impossibility to shield the identity of the research site (see ‘An Invitation to a Research Journey’), I should probably have warned them about the possible consequences of the research, especially in their personal regard (see Hammersley & Atkinson, 1995). As they held specific positions within the Junior

College at the time of the research, I could not even protect their identity by using pseudonyms. Thus, it was also to redress my oversight in their regards and their probable ‘naivety with qualitative research’ (see Smith, 1990; cited in Zeni, 1998) that I introduced the ‘dummy dates system’ to which I referred in the ‘Invitation’.

5.2.2 Selecting Participants and Building Trust

Beside myself, there were twelve other PMI teachers during Year 2 of the study. Seeing this as a manageable number, I decided to invite all twelve to take part. My prime consideration was that, within the available time and resources, I could still engage with all of them in a wide range of valuable experiences that are typical of information rich cases like mine (Patton, 1990). The only restriction I imposed on myself was not to add new participants in subsequent years. I was determined that, barring research dropouts or the emergence of what Patton (1990) calls ‘cases of special interest’, I would continue with my original participants.⁶

My reluctance to increase the number of participants can also be seen in the light of my desire to focus on a fixed number of teachers with whom I, as a researcher rather than as a colleague, could establish trust, rapport and authentic communication. In the knowledge that trust, even between professionals, has to be created (Simons, 1982), I set out right from the start to earn as a researcher my participants’ trust whenever and wherever I could. My fieldwork journal is indeed replete with episodes, both on and off the field (the latter due to my custom of socialising regularly with some colleagues outside school hours), that illustrate how hard I worked to establish this trust, even in the eyes of the non-participating colleagues. For instance, on the numerous occasions when some colleague or other made ‘friendly requests’ that I divulge information about what particular participants were ‘saying’ and ‘doing’, I always answered in a firm voice, audible to all present, that I would never do it. This stand, even if it did not stop the more assiduous colleagues from passing similar trust-breaching requests, sent the unequivocal message that I could be trusted.

⁶ As neither of these two eventualities materialised, when two other colleagues started teaching PMI classes in Year 3 of the study I did not ask them to join in. On the contrary, I decided to include myself as the ‘thirteenth participant’ when I became aware at one point in the study that this would give an added dimension to my data and interpretations (see sections 6.0 and 6.1).

5.2.3 Contacting the Participants

I decided to prepare my friends and/or colleagues for my 'new' role within the mathematics department by first creating an awareness about my study (Ely et al., 1991). By then, everyone knew that I was studying 'something'. This much they knew from my study trips to the UK during the two previous summers. Capitalising on this knowledge, I began 'informally' paving the way for my 'real' access negotiations by informing colleagues, whenever they would kindly query about my studies, that since I was particularly interested in the PMI option I would soon be approaching the PMI teachers within our department to invite their participation in my research. In a department of twenty teachers who share four offices in the same corridor, I felt certain that in a matter of days the word would spread around. In fact, once I started the proper access negotiations, most PMI colleagues admitted to having already heard about my research and appeared willing to listen.

Probably because I knew that my direct demands on the principal, registrar and subject coordinator would only be minimal, I had not felt uneasy about contacting them. But approaching my colleagues – the 'real' negotiation stage – was a different story, as I knew that I would be asking them for a very serious commitment. My uneasiness, as the following fieldwork journal entry shows, resulted primarily from having to approach colleagues with whom I previously had had little personal contact:

I have been dreading this moment for a long time. It's having to approach the 'older' generation that I'm finding harder. As I have no 'real' rapport with them, I have no clue how they will react. Sometimes I wish I had planned my study in such a way as to exclude their participation. (December, Year 2)

Wanting to appear in front of my colleagues as reassuring as possible so that I could put them at ease, earn their respect and eventually their collaboration, I decided to start approaching those colleagues, mostly friends, who had previously shown interest in my studies. This strategy permitted me to face the remaining colleagues, practically the more senior ones outside my social circle, when I had already gained some experience and confidence in facing and handling these negotiations. By the time I came round to what I had foreseen as potentially tougher talks, these PMI teachers were already

expecting me to drop by their office. Their expectancy and general good disposition in my regards permitted me to proceed more serenely with my requests. At the end of the day, even though none of my colleagues was over enthusiastic about participating, all consented (either immediately or after thinking things over a couple of days) to participate. I realised later that had I decided, as I was tempted to do, to exclude *a priori* some colleagues either for fear of rejection or due to my preconceived uneasiness about working with them, I would have lost important contributions to my understanding of TCAP.

5.2.4 Seeking Informed Consent

I figured out that only by coming up with reasonable demands and adequate guarantees could I hope for my colleagues' participation. Guided by the principle of informed consent, I tried to embody these demands and guarantees within a standard, formal access letter – much along the lines of what Seidman (1998) calls 'consent form' – that I wrote before approaching them individually. Apart from facilitating the growth of trust and helping to ensure its survival (Nias, 1981; cited in Simons, 1989), I hoped that this letter would send a clear message that everyone, irrespective of whether one is a friend or just a colleague, would be treated equally during the study. In particular, the letter referred to the purpose of the research, my independent researcher status, the participants' role in the research, and the measures intended to protect the interests of the participants. The letter informed my PMI colleagues that the purpose of the research was 'to explore your classroom experience with particular reference to issues related to assessment'. I was purposefully vague here for two reasons. First, I was then still at the stage of coming up with and evaluating different possible routes. Second, had it been possible to be more specific, I still had no wish to provide information that may influence the participants' behaviour in such a way as to invalidate the findings (Hammersley & Atkinson, 1995). On the contrary, I was very clear in the letter about the independent status of my research study. In particular, I wanted to reassure my colleagues that the study was not sponsored by, or in any way answerable to, people who have, directly or indirectly, power over them (Merriam, 1998).

With regards to their participation, the letter specified that I would be observing and recording what goes on in our department in relation to my research interests. It also informed them that I would be interviewing participants on a number of occasions, and that these interviews would be tape-recorded so that I could later on transcribe them myself. I assured them that no one, other than myself, would ever have access to their interview tapes or scripts. Finally, I reserved the right to use direct quotes in the write up from my observations and interviews. The letter also dealt with a number of ethical considerations. In particular, I bound myself to confidentiality on all matters arising from the research process, the use of pseudonyms, their right to review own transcripts and to affect changes, and their right to withdraw from the research at any point. The letter also pointed out that it would not be possible to shield the identity of our school. To counterbalance this I offered two guarantees. First, the thesis would not be available for general perusal in Malta. Second, should publications follow, the participants would have the prepublication right to edit and censor parts in order to protect their identity.

Although one-to-one access talks are singular affairs, I tried to somewhat standardise them by embedding the following elements in each contact: (i) approaching teachers only when they were alone in their shared offices and during periods when we were unlikely to be interrupted; (ii) presenting the access letter, and explaining and elaborating on its contents; (iii) encouraging teachers to react openly to my proposals and dealing with their queries; (iv) insisting that I am a learner attempting to see life through their eyes (Ely et al., 1991) and that it is therefore not my intention to judge or evaluate what teachers say and do; and (v) inviting formally the teachers to participate in the study and giving them ample time to reflect on my proposal. I was moreover careful, both in what I wrote in the letter and in my subsequent verbal dealings, to include only promises that I could keep. At the same time, though, I pragmatically chose to follow Foster's (1989) advice to stress during this initial negotiating stage only the less threatening aspects of my study. My intention was never deliberately to deceive. I simply did not want to scare off colleagues by mentioning the possibility that I might need to access their classes and make contact with their students at some future point. I preferred to leave this more delicate form of access, should the need arise, to when my field relationships with the participants were already established

(Hammersley & Atkinson, 1995). In fact, when I knew by January of Year 3 that my study would benefit from such evidence, I tried to renegotiate access.

5.2.5 Renegotiating Access with Participants

After spending the first term of Year 3 of the study at the University of Nottingham to finalise the research design and tools, I came back to the Junior College as a researcher only in early January (I only resumed my normal teaching duties as from April, Year 3). As I renewed my contacts, I stressed once again that their participation is voluntary. Everyone's positive response to continue participating led me to address the next item on my agenda – the renegotiation of access conditions.

My renegotiations, during which I again experienced uneasiness, practically followed the same procedures and reiterated the same guarantees of the original negotiations. As before, I included my requests in an access letter. Basically, I asked participants to allow me inside their classes as a non-participant observer for the duration of one week (i.e., for three lessons). I explained that this 'inside view' of their classroom would help me to better relate my forthcoming interviews with their reality and to understand more what they make out of it. Moreover, I specified that I would like to sit at the back of the classroom, take hand-written notes about the things I observe, and tape-record the lessons as an *aide memoir*. In my first few meetings I also sought their permission to speak to students. I justified this request by saying that this would help me understand how students experience the PMI option, especially in matters related to assessment. But once it became apparent that most teachers were evidently 'unsettled' by my request to contact students, I decided to drop it altogether in the remaining renegotiations. I did not cherish risking my chances to enter classrooms because of it. Consequently, I removed reference to this particular request in the amended version of the second access letter that all participants, except for one, eventually endorsed. This participant motivated the decision not to be observed in class by acute feelings of embarrassment.

My failure to access students meant that my understanding of TCAP depended exclusively on teachers' practices and perspectives. My data was thus limited to what

the participating teachers were 'willing to share' with me. Whilst I understood that this is far from ideal, I still had to sacrifice gaining knowledge from students for fear that this eventuality could dent my rapport with the teachers – who were, after all, my main research focus. Much to my surprise, I could sense that the teachers generally saw my request to access students as an attempt 'to pry behind their back'. For some reason or other, and in spite of my assurances to the contrary, some seemed moreover convinced that by acceding to my request they would be exposing themselves to 'student attack' without the possibility of replying to 'rectify the record' (as I had specified that whatever students say would be treated as confidential). Hindsight suggests that their resistance might have arisen from my insistence during the initial negotiations that the study focussed on their 'classroom experience' (see section 5.2.4). The teachers probably interpreted this to mean that whatever data I collected must be 'about them'.

5.3 The Fieldwork Phase

The fieldwork phase refers to the two scholastic years (i.e., Year 2 and Year 3) during which my role as researcher was overt.

5.3.1 The Fieldwork Journal

In the field, both during the covert and overt phases, I kept detailed field notes. This was an account of what I had seen and heard at school, without interpretation (Maykut & Morehouse, 1994). Field notes entries range from the voluminous interview transcripts to some minor field episodes that I recorded in a couple of sentences. Whilst I tape-recorded and later transcribed the formal interviews, most of the remaining field notes were initially short handwritten notes recorded discreetly at school at the first available opportunity in order not to render the participants unnecessarily self-conscious and uneasy about my 'research' presence. But later that same day, I habitually developed at home these sketchy pictures into detailed accounts. The idea here was to capture the substance of a conversation (or observation) rather than a flawless verbatim reproduction (Bogdan, 1972; cited in Maykut & Morehouse, 1994). The only occasions during which I took notes in the presence of the participants were when I observed their classes (which I also tape-recorded).

The fieldwork journal, in addition to the factual descriptions of the field notes, includes my personal commentary. I was very careful to distinguish in the journal between observed and inferred behaviour, between what I sensed (field notes) and what sense I made of it (commentary) (Wolcott, 1994). The comments in the journal represent my feelings, reactions, initial interpretations, and even working hypotheses. They are my thoughts about the setting, people, and activities, and include ‘relations to myself’ in the process of the research (see Griffiths, 1998). Through these writings, I was also engaging in what Merriam (1998) calls ‘some preliminary data analysis’.

5.3.2 The Interviews

Throughout the study, I conducted three semistructured interviews with each participant. The first interview, what I call the ‘background’ interview, took place during January of Year 2. A year later, between January and March of Year 3, I first observed each participant in class for a week (i.e., three lessons) and soon afterwards conducted the second and third interviews, what I call respectively the ‘mathematics’ and ‘assessment’ interviews. This second phase of interviewing (and classroom observations) was however preceded by a pilot-test with a PMI teacher from another school. This served to practise my interview guides (see section 4.3.3.1) and also to try out the procedures planned to record the data from classroom observations.

5.3.2.1 Preparing the Interviews

For a start, I relied on my readings and the numerous leads that kept emerging from the field to formulate the interview questions. But it was my ongoing reflections in the fieldwork journal that proved to be the main inspiration behind the development of the three interview guides (Ely et al., 1991). As I prepared for each interview, I repeatedly reviewed the arising questions with the help of non-school friends and mentors – a phase in which the worth of each question was evaluated. By the end of this exercise – which saw many questions being rejected and others introduced – my interview guides contained questions that could be classified along three of the six types of interview questions outlined by Patton (1990). These were ‘experience/behaviour’ questions, ‘opinion/values’ questions, and ‘background/demographic’ questions.

In particular, I used the practice-related experience/behaviour questions to elicit descriptions of experiences, behaviours, actions, and activities that would have been observable had I been present. The limited nature of my classroom observations heightened further the importance of their inclusion. On the other hand, I used opinion/values questions, which concern the cognitive and interpretive processes of people, to explore what the participants think about the issues involved. Finally, I used background/ demographic questions to fill in the gaps in my prior knowledge of the participants. They were particularly useful in providing indications about age, education and training, and teaching experience. As part of my preparation, I also got well acquainted with the interview guides so that I could confront interviews with confidence and approximate a very natural conversational style.

5.3.2.2 The Purposes of the Interviews

The ‘background’ interview was rather general and exploratory in nature. It was an opportunity to converse with participants about their school experience (particularly within the mathematics department) and the matriculation system (particularly the PMI option). I was moreover able at various points in the interview to unobtrusively gain information about the participants’ professional experiences and demographic data (Maykut & Morehouse, 1994). These initial interviews also provided an insightful ‘first peep’ inside PMI classrooms. But it was the second interviewing phase (in conjunction with classroom observations) that provided the bulk of my data. With each participant (except for the teacher who did not wish to be observed – see section 5.2.5), I began by observing his or her PMI class for a whole week (i.e., three lessons). Soon after this, a couple of days at most, I first conducted the ‘mathematics’ interview and then the ‘assessment’ interview. This interviewing order made it possible to explore the extent to which, if at all, teachers relate the teaching and learning of mathematics to assessment before I could directly address the issue during the ‘assessment’ interview. By observing PMI lessons before these two interviews, albeit for a limited period, I largely familiarised myself with what Silverman (1993) calls ‘the routine’ and obtained rich illustrations of classroom practice. I was thus in a position to acquire interview meanings from observations (Dexter, 1970; cited in Hammersley & Atkinson, 1995) and to embed our discussions in concrete classroom episodes.

During the ‘mathematics’ interview I explored teachers’ views about the teaching, learning and the nature of mathematics, and how they translate these into their PMI classroom practice. On the other hand, the ‘assessment’ interview dealt with teachers’ views on assessment, especially classroom assessment, and their PMI assessment practices. In both interviews, I used ‘projection techniques’ (see Patton, 1990). In fact, whenever I felt the need during the ‘mathematics’ interview to re-direct or deepen our conversation, I judiciously asked interviewees to react to one of the fifteen contrasting pairs of statements that I had prepared on separate cards (see Appendix I). But whilst I never made use of all, or even most of, the cards in any one ‘mathematics’ interview, I invariably sought the interviewees’ reaction to the four chosen tasks (see Appendix II) during the ‘assessment’ interviews. As one moves from Task 1 to Task 4, the nature of the task evolves along a continuum from the traditional examination type to the non-traditional type. In particular, Task 1 and Task 2 may be described as ‘routine’ or ‘traditional examination type’ tasks, with Task 1 being close-ended and highly scaffolded, and Task 2 being similarly close-ended but less highly scaffolded. On the other hand, Task 3 and Task 4 are ‘nonroutine’ tasks. Whilst both are unscaffolded, Task 3 is close-ended and Task 4 is open-ended. In addition, Task 1 and Task 2 may be categorised as lower order tasks as they only require lower level thinking from students, whilst Task 3 and Task 4 are higher order tasks as they require higher level thinking from students.

5.3.2.3 Conducting the Interviews

For the interviews, which I conducted in a quiet and private room at school, I wore my ‘usual’ school clothes. To further stress a collegial atmosphere, I decided in favour of a side-by-side seating arrangement next to a table. My efforts to build a non-hierarchical interviewing relationship were however not always successful. For although I regarded and treated ‘all’ participants as authoritative professionals with whom I desired to engage in serious professional discussion, some interviewees at times still cast me in the role of ‘expert’, usually citing either my ‘more advanced’ studies or their inexperience to justify this. During the interviews, which lasted between 45 and 75 minutes each, I equipped myself with the interview guide, a small good-quality tape recorder (which I placed in front of us on the table), recording cassettes, and spare

batteries. In addition, I took for the ‘mathematics’ and ‘assessment’ interviews the ‘projection techniques’ and teachers’ class records and examination predictions. I began each interview by reminding the participant about the purpose of the interview. Then, I gave the interviewee an opportunity to ‘organise his or her head’ by first asking general questions leading up to matters that I wanted to discuss in detail later on (Lincoln & Guba, 1985). From then onwards, the questioning sequence varied from one interview to the next, depending on the responses I was getting. My interviewing style was characterised by the use of primarily open-ended questions and follow-up probes, an easy rhythm facilitated by ‘interrupting gracefully’ when no new material was forthcoming, and a healthy ‘talk turn’ with respondents.

5.3.2.4 Transcribing the Interviews

I chose to transcribe the interviews myself for two main reasons. First, this guaranteed participants the confidentiality negotiated during the access talks. Second, I saw this as an opportunity to increase my familiarity with the growing body of data (Ely et al., 1991). I wrote an almost *verbatim* transcription of the recorded interviews,⁷ restricting myself to limited and purposeful editing in order to facilitate the communication of meaning of the participants’ stories. Without this editing, statements that were perfectly coherent within the context of a living conversation would have become incoherent or even contradictory. Although I should have ideally transcribed interviews whilst they were still fresh in my mind (Maykut & Morehouse, 1994), this was not always possible in my study. Sometimes, especially during the second interviewing phase, I could only start transcribing an interview weeks after I had recorded it owing to the sheer number of interviews and classroom observations involved. To partially make up for this, I referred to my post-interview reflections in the fieldwork journal before actually proceeding with the transcription. This helped me to re-remember memories and to re-immersify myself in the interview episode. Giving the transcripts back to interviewees for review produced, at best, only minimal changes – which I saw as an indication of their overwhelming acceptance of my rendition. I did however use this

⁷ Although I basically conducted the interviews in Maltese, both the interviewees and myself repeatedly code-switched between Maltese and English in the process. But whilst my actual transcription faithfully reflects this local phenomenon (see A. Camilleri, 1995), I later translated into ‘full’ English those quotes that appear in the text, as I did with all other quotes taken from field notes.

review stage to ask participants to clarify, either in writing or verbally, certain parts of their transcripts. This permitted me to share and develop with them the zone of possible meanings of the original interviews (Kvale, 1996).

5.3.3 The Classroom Observations

On my first visit to each classroom, I informed the students that I would be observing and tape-recording some lessons, in addition to taking down notes. During lessons, I sat unobtrusively at the back of the classroom in order to minimise the teacher and students' reactivity to my presence (Maykut & Morehouse, 1994). Knowing that it is psychologically and physiologically impossible to observe everything that is going on in the classroom (McNamara, 1980), I joined lessons with a relatively open mind, only gradually directing my attention to what I saw as interesting or significant themes. I had initially feared that as a Junior College PMI teacher myself, I would be 'too familiar' with the PMI classroom context not to develop what Everhart (1977; cited in Hockey, 1993) calls a 'nothing happened syndrome'. But once I began observing the PMI world from this different perspective, I soon realised that every lesson was adding something to my growing understanding of PMI teachers' classroom practices.

I recorded my classroom observations on a form that I had developed during pilot testing. Apart from the 'unplanned' aspects of classroom life that caught my attention, this form also included: (i) details of teacher, class, lesson topic, room allocation, date and time, and visit number; (ii) number of students present; (iii) teacher's action zone, his/her interaction with students, work given and resources used; (iv) classroom layout plan; and (v) a temporal chart of events. Although I can never be sure of how influential my presence in class was (Hammersley & Atkinson, 1995), I have reasons to believe that reactivity to it over the three visits was minimal. For instance, all the teachers who had confided during access talks that my presence would make them feel uneasy, reassured me afterwards that by my second visit this uneasiness had either greatly subsided or ended completely. Moreover, most teachers reported that as my visits progressed, their students appeared less conscious of my presence and acted more naturally (e.g., less well-behaved; not afraid to ask questions; etc.). Even I noticed how the formal atmosphere in class triggered by my presence (Merriam, 1998)

began to ease off with time. This became evident when both teachers and students manifestly appeared less concerned to present themselves ‘only’ in favourable light.

5.3.4 Leaving the Field

During Year 4 of the study, when my research interests practically shifted from the field to ‘formal’ data analysis and writing at home, I started feeling and acting at school more as a teacher than as a researcher. Thus, even though I never physically left the field, my researcher *persona* still left eventually – a departure that began from my gradual state of mental relaxation in the field. I began noting this towards the end of Year 3, soon after I had finished observing classes and interviewing teachers. From that point onwards I could barely keep the previous ‘distance’ from the participants – a ploy I had adopted during fieldwork in order not to influence unduly the data sources. And as I began surrendering myself more frequently to the setting or to the moment (Hammersley & Atkinson, 1995), I had less to write and reflect upon in my fieldwork journal. But my continued presence in the field permitted me at least to ‘keep the door open to new data’ (see Ely et al., 1991). Nevertheless, when it became evident by mid-Year 4 that I was no longer coming in contact with fresh data, I willingly gave in to my urge to re-establish myself as a ‘teacher only’ within the department.

5.3.5 Reflections on my Insider Status

My ‘real’ participation before and during the study was a source of opportunity and advantage over less immersed roles (Pollard, 1985). Not only did the research problem emerge from the field, but I also entered the field already knowing the people who could provide access and information, what constitutes appropriate conduct and demeanour at the site, and the roles and relationships amongst the participants. This facilitated access and made it easier for me to establish and maintain a good, trusting relationship with the participants. As a ‘real’ insider, I could also use familiar language that favoured ‘sharedness of meanings’ (Fontana & Frey, 1994) and better achieve invisibility in the field (Hockey, 1993). Again, had I not been a teacher at the site, it would not have been possible to adopt – certainly not for the extended period that I did – a research methodology that draws on ethnography. The making sense process of the

resulting rich data was facilitated in turn by my 'localised' broad cultural and common sense knowledge. But although I accumulated as an insider knowledge and understanding that could not have come in any other way, there were however a number of methodological and personal constraints attached to this privileged position.

I could not, for instance, easily adopt the 'novice role' (see Hammersley & Atkinson, 1995). My attempts to 'feign naïve', when I either wanted to explore participants' meanings or to 'put on record' something that they expected me to know, were largely met with polite signals of impatience and incomprehension. Moreover, I could have possibly been denied access to students, at least in part, because I was an insider. Some participants in fact told me during renegotiations that they do not like their students 'to talk about them', but even more so when the person receiving the information is a colleague. On the other hand, believing that to gain deep insights about core meanings and experiences I had to create some 'distance' from the social situations and relationships that I studied (Adler & Adler, 1994; Gans, 1982; Hammersley & Atkinson, 1995), I constantly fought back the natural comfortable sense of being 'at home' within my own school. To mentally distance myself from the 'too familiar' context, I used reflective writing and generally kept my thoughts to myself at school. This 'holding back', in particular, made me however feel guilty of neglecting my duties within the department. My discomfort grew more intense when I held back on matters about which I felt strongly and could have made a valid contribution. Another consequence of my marginality was the 'great sense of loneliness' that I often felt during the study. For in addition to spending most of my time outside school working in isolation at home, I also had to endure a mental form of isolation at school. Being an insider, I had also much to lose on a personal level had things gone wrong. For I was not only accountable to participants, but also to colleagues and friends.

5.4 Turning Data into Research Findings

5.4.1 Analysing and Interpreting the Data

My fieldwork journal served to produce in its totality what Bogdan and Biklen (1982) call the 'data about the data'. It was this process, which is based on "conversations

with oneself about what has occurred in the research process, what has been learned, the insights this provides, and the leads that suggest future action” (Ely et al., 1991, p. 80), that helped me move forward the methodology and data analysis, and ultimately my understanding. The challenge was to hold different interpretations of the unfolding evidence until deciding which interpretation is merited by the data being collected (Maykut & Morehouse, 1994). In an effort to allow the meanings and perceptions of the participants to emerge, apart from submitting their interview transcripts for review (see section 5.3.2.4), I initially also asked participants to respond to preliminary patterns and themes as they developed from my ongoing reflections on the data (Maykut & Morehouse, 1994). Had this worked, not only would I have been able to check the interpretations made, but I would also have generated further insights (Broadfoot, 1996). As it turned out, however, the participants’ interest was largely limited to trying to identify themselves and others in my writings. Recognising that this may be seen, at least by some, as a breach of my confidentiality guarantees, I subsequently decided to stop this practice at the cost of weakening my search for multiple realities. The complex, reflective process of moving from the field to the text to the reader (Denzin, 1994) has consequently been my sole responsibility.

As an insider, I found it very fascinating and rewarding to identify patterns in the data and to hesitantly, step-by-step, attempt to construct a deeper understanding of the events in which I participated daily. This was particularly the case during the data analysis proper – that is, when I finally set out, after sensing that I had tapped all the available evidence, to systematically organise my data and to start making conclusive, encompassing sense out of it. My way of seeing developed through Boyatzis’ (1998) three phases of ‘thematic analysis’. These include recognising an important moment (i.e., seeing), which precedes encoding it (i.e., seeing it as something), which in turn precedes interpretation. In practice, I began this thematic analysis by reviewing (through reading in most cases) and reflecting upon the various data sources collected during the study. As I immersed myself within the data, literally months on end, I began coding it into what I am calling ‘units of meaning’ – a process during which I attached my initial isolated meanings and understandings to interview transcript phrases and sentences, other field notes and documentary sources. As the units of meaning began to pile up from this cocktail of evidential sources, I gradually started to

notice new links and meanings within and amongst the different data sources. As a result, localised meanings began to evolve into more general themes. These new and deeper understandings eventually led to my data interpretation proper. What I can present, though, is ‘uncertain’ knowledge bearing my marks that is self-consciously situated in its context and always subject to revision (Griffiths, 1998).

5.4.2 Key to Interpreting the Research Findings

A case study can communicate enduring ‘truths’ about the human condition through the portrayal of a single instance locked in time and circumstances. The theory-building potential of my study thus lies in the fact that although it focuses on a small group of teachers within a particular context, its findings and conclusions still have the value to refine theory and suggest complexities for further investigation, as well as to help establish the limits of generalisability (Stake, 1994). In recognition that the purpose of my study is to represent the case, not the world (Stake, 1994), the only claim I make is that my findings and conclusions represent what Bassey (1999) calls ‘fuzzy generalisations’. The fuzzy generalisation, which is a qualitative measure based on the study of singularities, “typically claims that *it is possible, or likely, or unlikely that* what was found in the singularity will be found in similar situations elsewhere” (Bassey, 1999, p. 12; emphasis in original). This embedded ‘element of uncertainty’, far from being an admission of frailty in the way I conducted the research, reflects instead the many variables that underline the complexity of human behaviour and interaction. What I find particularly exciting about a fuzzy generalisation is that (whereas with a statistical generalisation the reader is only expected to understand and absorb) it actually invites the reader to take a more active role. The reader is invited to enter into discourse about it by reading the evidence that supports it, by discussing it with people engaged in similar situations, by reflecting on the issue, and even by testing it out in his or her own environment (Bassey, 1999). Research subsequently becomes cumulative when the fuzzy generalisations drawn are amended according to the new evidence that surfaces when the study is replicated in different environments and circumstances.

Part FOUR

THE DATA ANALYSIS

CHAPTER 6

Meeting Teachers and Entering Classrooms

6.0 Deciding to Become One of the Participants

When I decided to invite my twelve PMI colleagues to participate in the study (see section 5.2.2) I was still seeing myself in the role of ‘researcher only’ in spite of being a PMI teacher myself. But once the data gathering process proper was set into motion, I found myself increasingly being drawn into comparing and contrasting my own assessment beliefs and practices with those of the participants. At first, it was largely a matter of using the stories of these significant others to engage in both personal and professional reflection. However, I gradually realised that my inclusion as the ‘thirteenth participant’ in the study would further enrich the data and thus contribute to my understanding of TCAP. This understanding was based on three main reasons. First, I reasoned that my experience as a teacher is arguably as relevant to understanding TCAP as that of my colleagues, for we all work within the same department, teach the same option, and basically operate within similar circumstances and constraints. Second, once I realised the extent to which my educational background and interests differ from those of my participating colleagues, I began to appreciate the comparative research possibilities that this raises. Third, I figured out that my continuing attempts as a teacher to turn alternative assessment ideas into classroom practice – apart from being an ongoing learning experience in itself (see Black et al., 2003) – would offer a unique perspective on the process and prospects of bringing assessment change at classroom level.

This ‘late’ decision to include my participation in the study led to some noteworthy developments in the research process. On the human level, once I informed my colleagues that I was ‘a participant just like them’, I could sense that a closer research relationship began to develop between us. In particular, although we still continued to relate mostly to one another ‘in the field’ as researcher and researched, there were numerous occasions during which our interactions, more than anything else, took the semblance of exchanges between participants. My participation thus made it somewhat easier for the other participants to understand something I had tried to

clarify during the access negotiations (see section 5.2.4) – namely, that ‘Michael the researcher’ was interested in understanding ‘us’ and ‘our practices’ rather than in evaluating ‘them’ and ‘their practices’. Another important development following my inclusion concerned the gathering of data. Apart from adding my classroom documents (see section 4.3.1) to those of the other participants, I saw fit to substitute the data I gathered from their interviews and observations with pertinent reflective writings in the fieldwork journal.

6.1 Introducing the Teachers

The twelve PMI participating teachers and myself are listed below in alphabetical order. First names are used, as is the custom of calling each other within our department. The information given includes indications of age, path to teaching career, teaching experience, teacher training, academic qualifications, office allocation (see sections 2.5.1 and 2.5.2), and the PMI class level (i.e., first or second year) taught during the research.

- Fifty-year-old **Andrew** holds a BSc degree in mathematics and has been teaching for nearly thirty years. Other than the first few years he spent in a secondary school, he has since always taught at sixth form level. Even though Andrew had always desired to become a teacher, he has never had any teacher training. Andrew, who joined the mathematics department with the first group of teachers in 1995, is located in office C and teaches first year PMI classes. Apart from giving private lessons, he also teaches mathematics at undergraduate level on a part-time basis.
- **Angelo**, who is in his early fifties, is a BSc graduate in mathematics. He has been teaching for about thirty years, and his teaching experience can be evenly divided into two parts – the first half at secondary level and the second half at sixth form level. Angelo chose teaching – a decision he now regrets – as it was then a relatively secure and well-paid job. He joined our school when it was set up in 1995. Angelo, who lacks teacher training, is in office C and teaches first year PMI classes.

- Thirty-year-old **Carmel** was ‘encouraged’ into teaching by his parents who convinced him that it was the ideal job for him in view of the ‘holidays’ – a career decision that he is now beginning to question. He has a BSc degree in mathematics and a PGCE. Carmel joined our school when it was set up in 1995. Prior to that, his teaching experience was limited to just one year in another sixth form college. He is in office B and teaches second year PMI classes.
- Thirty-year-old **Jackie** – who ‘still’ has no regrets for becoming a teacher – joined the profession as it appeared to be the only career opportunity available in Malta for someone specialised in mathematics. After finishing her BSc degree in mathematics, she joined the PGCE course and is currently studying part-time for an MSc degree in mathematics. Before joining our school in 1995, Jackie had taught for one year in another sixth form college. She is in office B and teaches first year PMI classes.
- After finishing her BSc degree in mathematics, **Kathleen** decided to join the PGCE course so that she would have further career openings in the future. During her first teaching practice she decided however to make teaching her career. Thirty-year-old Kathleen had no teaching experience when she joined our department a couple of years after the school opened its doors in 1995. She is in office D and teaches second year PMI classes.
- Fifty-five-year-old **Mario**, who has been teaching for more than thirty years, decided to become a teacher so that he could earn a scholarship in mathematics. Although he initially taught for a short time at primary and secondary levels, he has now been teaching at sixth form level for more than twenty-five years. Mario, who joined our school in 1995, holds BSc and MSc degrees in mathematics and has a Teacher’s Certificate. He is in office C and teaches first year PMI classes.
- **Matthew**, who is in his late forties, has been teaching for more than twenty years. He initially considered teaching – a career move he now regrets – a natural choice for someone who likes mathematics. Except for the initial couple of years, when he taught in a secondary school, he has always taught at sixth form level. Matthew also

teaches mathematics at undergraduate level on a part-time basis and gives private lessons. He holds BSc and MSc degrees in mathematics and has a PGCE. Matthew, who joined our school in 1995, is in office C and teaches first year PMI classes.

- **Michael:** It was never my intention to become a teacher. I just wanted to continue my studies in mathematics at degree level. However, as the University of Malta was by that time no longer offering BSc courses, my only option in 1980 was to join begrudgingly the BEd course with specialisation in mathematics. My attitude to teaching nevertheless changed completely during my first teaching practice when the smiling faces, eagerness, affection and loyalty of my seven-year-old pupils won me over. I have now been enjoying teaching – with experiences spanning from the primary to undergraduate levels – for more than twenty years. My ongoing commitment to teaching and keen interest in educational issues in general have led me over the years to invest in my professional development (reading for my MEd and PhD degrees are but the formal side of this coin) and to participate actively in a number of educational initiatives (see section 1.0). At school, which I joined in 1995, I am in office B and teach second year PMI classes.
- Forty-year-old **Nicholas** always wanted to become a teacher. He holds a BEd degree with specialisation in mathematics (and experience that he still ‘despises’, having only joined due to the lack of BSc courses at that time) and an MSc degree in mathematics. Nicholas taught briefly at primary and secondary levels before moving to sixth form level some fifteen years ago. Besides teaching at our school, which he joined in 1995, he gives private lessons up to undergraduate level. Nicholas is in office B and teaches first year PMI classes.
- Teaching has been **Ray**’s much-desired career choice since he was a little boy. Being now almost sixty, he has already been teaching for more than forty years. After teaching mathematics at secondary level, he moved to sixth form level some twenty years ago. Ray, who also gives private lessons, has a Teacher’s Certificate and holds a BSc degree in mathematics. He joined our school when it opened in 1995. Ray is in office C and teaches second year PMI classes.

- **Renzo**, who is in his late forties, has been teaching for more than twenty years. His main motivation for becoming a teacher – a decision he does not regret – was the lack of job opportunities at the time. He taught mathematics in secondary schools before moving to sixth form level some fifteen years ago. He holds a BSc degree in mathematics and has a PGCE. Renzo, who joined our school when it opened in 1995, is in office B and teaches second year PMI classes.
- **Rita** always had the desire to become a teacher. Her first teaching experience was at our school, which she joined a couple of years after it opened. Rita, who is in her late twenties, has BSc and MSc degrees in mathematics. Except for the year she spent in the BEd course before changing university faculty from Education to Science (she initially went for ‘education’ as she had feared the mathematics content of the BSc course), Rita has no teacher training. She also conducts mathematics tutorials at undergraduate level. Rita is in office D and teaches second year PMI classes.
- Fifty-five-year-old **Stephen**, who joined our school when it was set up, has been teaching for more than thirty years. Although he enjoys teaching, he actually decided to become a teacher as it gave him the opportunity to continue with his studies in mathematics. Except for the first few years when he taught in a secondary school, he has always taught at sixth form level. Stephen holds BSc and MSc degrees in mathematics, but has no teacher training. He is in office C and teaches first year PMI classes.

6.2 The Rationale behind Michael’s Story

Amongst my closest colleagues within the mathematics department, I have increasingly earned the reputation of being ‘an education man’. This nomenclature reflects an acknowledgement on their part that my educational background and interests contrast widely with those of most teachers within our department. Being one of only three teachers in the department to have a BEd degree (Nicholas is one of the other two) and one of only two teachers to have an MEd degree (the other one did not participate in the study) already sets me aside from most of my colleagues who have their degrees more intimately linked to mathematics content – basically BSc and MSc

degrees in mathematics. But as my interactions with the other participants revealed, it is not just a matter of qualifications. The crucial point is probably the different kinds of sensitivities on our parts. By their own admission, my participating colleagues appear to lack a professional life (see Schoenfeld, 1999). Irrespective of whether they are teacher trained or not, their commitment to teaching largely begins and ends with preparing and delivering lessons, and ‘keeping up to date’ with the MATSEC examination papers for which they prepare the students. Their ‘restricted’ professional lives – with some claiming to be too busy outside school (‘giving private lessons’ and ‘studying’ were the two most frequently mentioned activities) to engage in time consuming professional activities – contrast sharply with my ongoing efforts to keep abreast with educational issues and developments, both locally and abroad. I consequently possess an educational baggage that the other participants, to a large extent, do not share.

Thus, even though I decided to become one of the participants, the fact however remained that my form of participation was unique – a uniqueness that went beyond being the only participant being researched by him or herself. Apart from having educational assessment as my area of study, what particularly distinguished me from the rest was my awareness and understanding of local educational policies and initiatives concerning the need for greater emphasis on formative assessment at classroom level, the rationale of curriculum breadth at sixth form level, and the ongoing efforts aimed at providing quality education for all whilst continuing to make post-secondary and tertiary education more accessible (see section 2.6). My inclusion consequently adds the participation of a well-informed teacher who is trying to improve his classroom assessment practices within the same institutional, curricular and societal dynamics faced by the other participants. To reflect this distinctive position, I chose to present my data inside the ‘bracketed’ segments entitled *Michael’s Story*. This ongoing construction of my own case study – which continues throughout the entire presentation of the results – may be viewed as an ‘open window’ on my classroom reality and way of seeing things. Besides mapping out further the positions from which I draw my research understandings, comments and conclusions, *Michael’s Story* extends the research data to explore TCAP within a scenario that is otherwise unfamiliar to this study.

6.3 Getting to Know the Teachers

6.3.1 Their Views on Teacher Education

Andrew, Angelo, Stephen and Rita (who actually followed the first year of the BEd course) are the only teachers who have no teaching qualifications. But whilst all four of them never exhibited any signs of regret about their lack of training, comments made by teachers who followed either the BEd or the PGCE courses depict their teacher training experiences either in a very negative (e.g., Matthew: “*a wasted year*”; Jackie: “*an imposed and senseless thing*”; and Nicholas “*doing methodology was a waste of time*”) or, at best, in an unenthusiastic manner (e.g., Renzo: “*what I used to see as a waste of time may have provided me with some background*”; and Carmel: “*better than I had feared*”). In contrast, Ray and Mario, the two senior teachers who had attended the now defunct two-year course leading to the Teacher’s Certificate, spoke very positively about their training. But more than the professional or academic aspects, they both lauded mostly the social aspects of their experience. The teachers who had been previously employed for some time with the Education Division before joining our school reported that, albeit on very rare occasions, they had attended short in-service courses organised by the Division. They all concurred in describing these in-service training programmes as largely unyielding experiences.

When I asked the participants to order ‘teaching’, ‘learning’ and ‘assessment’ according to their classroom priorities, ‘assessment’ always ended up at the bottom of everyone’s list. This lack of focus on assessment – which is a common phenomenon amongst teachers (Calfee & Masuda, 1997) – has parallels with their assessment training. The vast majority of teachers maintained either to have had no such training, or else not to remember if they had or not. In fact, only Jackie, Kathleen and Carmel (i.e., the younger generation of PGCE trained teachers) mentioned some form of assessment training, even if very rudimentary. This is how Jackie described her assessment training:

Throughout the course we had just one lecture on assessment ... it was for all PGCE students, not just the mathematics ones. ... We were asked to mark a short language paragraph and a mathematics problem. Our marks on the problem were more clustered than on the language piece ... the lecturer wanted to show us

that assessing languages is more subjective than assessing mathematics ... that assessing mathematics is a more exact thing. But that's all we did!

These participants thus reflect the international trend characterised by teachers receiving virtually no training in assessment (see Stiggins, 1992) and being, in any case, out of date with developments in assessment (see Impara et al., 1993). Still, only Kathleen answered directly in the affirmative when I raised the issue of whether they presently felt the need for assessment training or retraining. Their remaining reactions can be classified under two groupings that roughly contain the same number of teachers. The first group of teachers held the view that although what they know about assessment already serves them well enough, they would ‘not really mind’ if they were to attend an assessment course as long as it ‘was on a voluntary basis’, ‘offered practical suggestions’, ‘was run by competent people’, and ‘did not force them to change their practices’. On the other hand, the teachers in the second group – who appear to believe more strongly than those in the first group that teachers acquire competence in assessment as they move from novices to experts (see Sadler, 1989) – concurred that they neither see the need for nor do they intend to participate in any assessment training or retraining programmes.

“

Michael's Story

I empathise with my colleagues when they speak negatively about their teacher training and glorify ‘teaching experience’ (see section 6.3.2) because I, for one, often felt as a BEd student that I was learning more about teaching from my own classroom practice than from lectures at university. My disenchantment then arose primarily from the ‘detached’ manner in which methodology was being presented to us. Theory was theory, and practice was practice – two different entities that hardly ever met to enlighten one another. But I realised later that it is possible to bridge this senseless gulf between research and practice (see Mitchell, 1999). Nowadays, I consider good quality initial and ongoing in-service teacher education programmes as the key to the realisation of a truly professional and committed teaching community at all teaching levels. These programmes need to tackle the ‘bread and butter’ issues of teaching in such a way that the relationship between theory and practice is interactive (see Elliott, 1993a). This makes it unacceptable, for instance, that I had no assessment training whatsoever during the BEd course, when assessment is so fundamental to the teaching and learning process. To all intents and purposes, even if I had brushed with the world of assessment during my MEd studies, I only really came into close contact with it when I embarked on this project.

”

6.3.2 Their Views on Teaching Experience

All the teachers continually referred to ‘experience’ when speaking about teaching. Teaching experience emerged in fact as a characteristic that is highly valued by all teachers, irrespective of their professional backgrounds and experiences, for successful teaching. Whilst beginning teachers yearn to gain classroom experience and see it as the vehicle through which they can move from novice to expert status (e.g., Rita: “*I’m sure that my experienced colleagues are better teachers than I am because of their experience*”), the more experienced ones find their classroom experience a reassuring factor and, occasionally, a matter of pride (e.g., Matthew: “*Teaching is about communicating with students ... a skill that is only acquired after long years of teaching like I have*”). Their comments depict teaching as a craft, something one learns through hands-on experience and as a direct consequence of that experience (see Day, 1993; also Fang, 1996). Some of them also spoke about the importance of acquiring the ‘recipes’ of experienced teachers, both from amongst their ex-teachers and colleagues, as part of accomplishing the trajectory from novice to expert (see Hargreaves, 1993). There was also a general agreement amongst them that once beginning teachers gradually develop their own teaching style, they are very likely to retain it for the rest of their career. Matthew articulated this very clearly:

When I was still finding my feet in teaching, I used to try out different things in class ... but once I found the teaching style that suited me, I’ve stuck with it since!

This attitude of treating their ‘recipe knowledge’ as unquestioned and unquestionable (see Schön, 1983) is in line with Calderhead’s (1987) observation that teachers’ practices are hard to change once established and adapted to the context.

“

Michael’s Story

I have no difficulty admitting that all my past teaching experiences are a great source of support to my daily teaching endeavours. After all, they represent the ‘knowledge-in-action’ that I gained over the years through my skilled action as a professional (see Schön, 1983). But then, unlike my colleagues, I also realise that to rely on experience, impulse and intuition alone is to deny the fundamental role that reflection and professional development can play in guiding and improving my classroom actions. I consequently choose to see my teaching as ‘work-in-progress’ – something that needs to be consistently monitored and

improved upon – rather than as a finalised product that does not occasion further reflection. My ongoing efforts aimed at personal professional development are thus not a sign of weakness on my part. They signal instead my resolve to fight against the well-known teacher characteristic of reluctance or inability to change (see Broadfoot, 1996).

”

6.3.3 Their Views on the School

Most teachers expressed disappointment that our school, in spite of its university affiliation and pre-university vocation (see section 2.4.1), does not practise selectivity and even accepts students with minimum entry qualifications (see section 2.4.2). Their main argument was that ‘our emphasis on numbers rather than quality’ puts our school at a disadvantage in relation to the other local sixth form colleges – certainly those with whom ‘we’re competing directly’ for MATSEC examination pass rates – that have highly selective entry criteria. A further general concern amongst the teachers is that our ‘inclusivity’ makes it difficult for us, if not impossible, to provide the quality education we purport to offer our students in preparation for university studies (see Junior College, 2000a). The teachers are not however against ‘our many weak students’ continuing with their formal education, as long as this is ‘elsewhere’ – that is, in institutions that are less academically oriented.

Some teachers also argued that whilst non-selectivity at the point of entry is imposed on us from outside, the school administration is however to blame for doing practically nothing to rectify this situation once the students are ‘in’. These teachers, who would like to see the school ‘weeding out’ students who in their opinion should not be there in the first place, attributed this school’s failure to two main reasons: (i) *the school promotion system* – which includes an element of continuous assessment (see section 2.4.4) with which most teachers agree in principle if not with the ‘too generous’ overall weighting it gets – permits weak and undeserving students to make it to second year; and (ii) *academic discipline* at school is reputed to be very weak, as no action is ever taken against students who persistently neglect their studies and perform at unacceptable levels (in extreme cases, the teachers would favour the holding back of the maintenance grant [see section 2.1.1] and/or the suspension or expulsion of students from school).

“

Michael's Story

Unlike most of my colleagues, I welcome the increasing number of students joining our school and aspiring for university education. This emerging reality – which mirrors a worldwide tendency (see Broadfoot & Black, 2004) – calls for a school system in which, unlike ours, issues such as ‘promotion’ and ‘academic discipline’ are developed as a means to support learning rather than to judge it. I agree with the participants that many of our students arrive with difficulties in what has been traditionally seen as ‘basic O-Level mathematics’, that many of them are being promoted undeservedly, and that the school administrators are conspicuous by their absence in the learning process. But I do not then agree that by not letting students enrol at our school or by making it hard for them to continue we would be doing anyone, ourselves included, any favours. The presence of such students is instead living proof of how much there is still to be done if we are to make mathematics work for all students. My point is that “To fail children in mathematics, or to let mathematics fail them, is to close off an important means of access to society’s resources” (Schoenfeld, 2002, p. 13). We should thus be aiming to provide high quality mathematics instruction for all students. But in order to do this there must be serious rethinking about what sixth form mathematics, particularly in view of the level adjustments at secondary level (see sections 2.2.1.1 and 6.3.5), can realistically hope to achieve.

”

6.3.4 Their Views on the Mathematics Department

Most of the participants described the mathematics teaching staff as a group of people largely fragmented along office lines – a perception that Andrew crystallised when he said, “*There are as many factions in our department as there are offices*”. In line with this, many teachers spoke of only having personal affiliations and friendships, which go well beyond the school environment, with some colleagues with whom they share the same office. And whilst there were hardly any reported interactions amongst colleagues from different offices, the reported interactions amongst same office colleagues were of a predominantly social nature, not professional. The professional interactions mentioned by the teachers included ‘talking about students’, ‘sharing information about syllabus coverage’, ‘reviewing each other’s class tests’, ‘recounting classroom anecdotes’, and ‘discussing solutions to particular mathematical problems’. Most teachers claimed that, apart from the knowledge gained from such interactions, they are completely in the dark about what the other teachers actually do inside their classrooms. The general picture within the department is thus one of teacher ‘individualism’ (see Hargreaves, 1994) that excludes the much beneficial rich

professional dialogue amongst teachers (see Gipps et al., 1995). The most notable exceptions to this prevailing reality of teachers working in almost complete isolation from each other – which in Malta, where a culture of collaboration within schools is still lacking (Bezzina & Camilleri, 2001), is one of the major features of teaching (Bezzina, 2001) – are Rita and Francesca (one of the non-participating teachers). According to Rita, they plan and do everything together. The other teachers, apart from disclosing to never having worked within a team, generally showed a strong preference to work on their own. However, many still conceded that cooperation amongst colleagues does have its benefits, especially for beginning teachers in order to get them started. The few who actually expressed some interest in working closely with colleagues, nevertheless stressed that whatever form such collaboration might take, it should never impinge on their role of autonomous decision-makers inside their own classrooms.

At the same time, many participants expressed resentment for not being involved in the running of the department. Their grievances depicted the subject coordinator – who is otherwise generally liked by most teachers – as the archetype of the ‘authoritarian head’ (see Ball, 1987; cited in Boaler, 1997). Their complaints included ‘end-of-first-year examination papers that reflect only his ideas’, ‘using staff meetings mainly to inform teachers about decisions already taken by him or the school administration’ and ‘shooting down ideas put forward by other teachers unless these conform to his’. Many teachers moreover criticised his ‘custom’ of convening staff meetings only very rarely, preferring instead to approach teachers either individually or in small groups – an exercise that Jackie likened to ‘whispering in the ears’. A few teachers even confided their fear that such ‘tactics’ not only do not foster dialogue amongst teachers, but possibly also lead to bickering and pique. Contrary to current practice, most of the participants expressed a desire to operate in a truly collegial environment – a department where teachers are actively and openly involved during the whole decision-making process, and where they are also valued and sought for what they can contribute besides teaching. The teachers thus seem to want collegiality at departmental level, but not at classroom level where they still prefer to jealously guard their autonomy (see Hargreaves, 1994).

“

Michael's Story

Our lack of involvement in the decision-making and implementation processes within the department has always bothered me. With a subject coordinator who likes to keep the reins tightly in his hands, our involvement is practically limited to teaching the assigned classes. Indeed, as far as I can recall, I only contribute otherwise to the department whenever the coordinator approaches me (as he does with a few others) to proofread and/or work out the pure mathematics end-of-first-year examination papers that he sets. At classroom level, although I am wholly committed to help transform our isolationist departmental practices into more collegial ones, I cannot yet claim to be working within a team of teachers. To this day, I practically still work in almost complete isolation from my colleagues – my professional interactions with them are also largely restricted to the ones I listed above. So much so that before I began collecting my research data, I was as completely in the dark about my colleagues' classroom practices as all the participants professed to be. Somehow, although I think that the coordinator is also to 'blame' for this reality – for albeit it is true that he does not prevent us from working together should we so choose, his administrative style is not conducive to a collegial environment – I would still lay the bigger share of the 'blame' on my colleagues and myself. I, for one, have generally sought collegiality in 'distant' matters concerning the department (an eventuality that the coordinator, in any case, continues to oppose), but have been less forthcoming in extending collegiality to matters more directly related to my teaching (and this in spite of believing that this is the way to go). The Assessment Issue Group, which I set up and coordinated, has so far been my 'boldest' attempt to fight the isolationist practices within our department (see 'My Journey as a Teacher Continues ...').

”

6.3.5 Their Views on the PMI Option

The teachers overwhelmingly concurred that the demands of the Matriculation Certificate, especially in view of the number of subjects involved, are beyond most students – at least, the 'not better than average' students who constitute, in their view, a very big majority.⁸ They claimed further that students, as a result of spending most school days rushing off from one lesson to the next without any break, have problems with finding time for personal contact hours (see section 2.4.1) and often arrive home too exhausted to do any serious school work or study. The teachers argued that the Matriculation Certificate regulations (which require students to study at least one science subject – see section 2.2.1.2) are 'forcing' many students to take the PMI

⁸ The number of students joining undergraduate courses at the University of Malta has however continued to increase following the Matriculation Certificate's first examination session in 1997 (Zarb Adami et al., 1999).

option without any inclination whatsoever, without enjoying it, and without any intention or desire to continue with the subject, in some form or other, at university. This positions most PMI students in direct contrast with their teachers who all claimed to ‘love’ mathematics (which everyone described as his or her all-time favourite subject in which he or she always did well) and to feel competent and knowledgeable in it (in fact, many frequently and willingly referred to their mathematics content qualifications, which reach MSc level with almost half of them, with a certain sense of pride). Ironically, many teachers expressed the same ‘reluctance’ with regards to the PMI option that they attributed to most of their students – revealing in the process what Pound (1998) sees as a commitment to the notion of academic excellence at sixth form level symbolised by the ‘golden standard’ of A-Levels. In fact, only Kathleen, Renzo and Carmel claimed to have no particular preference between teaching at A- and I-Levels. The rest, for reasons that had mainly to do with the PMI content level and students, made it clear that they would only be too pleased not to teach the PMI option when, and if, they are ever given the chance.⁹

The teachers used phrases such as ‘not much higher than O-Level’ and ‘just O-Level plus’ to describe the PMI content that, in their view, is extremely low in comparison to the traditional A-Level content. In spite of such evaluations, they concurred further that many students still find the PMI option rather difficult, if not outright beyond their ability. These comments mirror the common belief amongst the participants that today’s sixth form mathematics students are largely inferior to those of ‘even some years back’. They gave two reasons for this. First, that secondary level mathematics no longer respects the traditional layer-cake structure that guarantees each course to meet the prerequisites of the next (see Romberg & Kaput, 1999).¹⁰ Second, that increases in student numbers inevitably results in lowering of standards (see Wiliam, 1996). The teachers disclosed however that albeit the issue of ‘falling standards’ affects all their students, it is even more pronounced with PMI students, whom they described as largely incapable of working on their own, accustomed to work things out by using some formula or recipe instead of thinking, depending almost exclusively on teachers

⁹ Some time after I finished the data collection, logistical developments within the department made it possible for teachers to teach only at A-Level. Most of my twelve participating colleagues ‘grabbed’ this opportunity and have since stopped teaching PMI classes.

¹⁰ They were particularly critical of the syllabus content covered in Paper 1 and Paper 2B of the SEC mathematics examination (see section 2.2.1.1).

to perform even the simplest of mathematical tasks, and not particularly indicated to further their studies in mathematics, even if ‘only’ at I-Level.

On top of this, the teachers see the PMI students as largely making no genuine effort to improve matters, being only prepared instead to do the barest minimum possible. By and large, the PMI students (about whose personal background all the teachers conceded that, in most cases, they barely know anything) emerged as the kind of students that most of my participating colleagues would rather not teach, as they fail to demonstrate most of their ‘good student characteristics’. These characteristics – which have much to do with my colleagues’ adherence to the ‘bright-person’ model of teaching (see Sedlak, 1987; cited in Fang, 1996) – include ‘willingness to learn’, ‘intelligence’, ‘curiosity’, ‘paying attention in class’, ‘regular attendance’, ‘making judicious use of contact hours’, ‘doing the assigned work’, ‘good memory’, ‘ability to work on one’s own’, ‘participating during lessons’, and ‘aiming high for examination success’. A further obstacle noted by the teachers was the generally large PMI student groupings (see section 2.5.3). Although most of the teachers, particularly the more experienced ones, claimed that syllabus coverage is not a problem with three hours a week, everyone insisted that they cannot afford to engage in time consuming activities or tasks.

“

Michael’s Story

Although I consider the recent widening of the local sixth form curriculum as a positive development, I think that the Matriculation Certificate (with its six subjects to be assessed primarily through summative examinations at the end of the two-year course) may be too demanding on most students. At least, this is what I perceive on a daily basis from my students. At the same time, I see in the PMI option an opportunity for many more students to continue with their mathematics education beyond secondary schooling (see sections 2.3.2 and 2.5.3). In the knowledge that a ‘can’t do maths’ accusation attested by lack of qualifications in mathematics “locates you as unsuccessful, and lacking in intellectual capability; it locates you on the edge of the employment and labour market, as virtually unemployable” (Gates, 2002, pp. 213-214), these increasing numbers would become a further reason to celebrate should they lead to genuine learning opportunities. Although I accept that PMI students tend to be generally weak in mathematics (certainly weaker on average than the A-Level ones – but then, the PMI option is a lot less demanding), they are otherwise not that different from the other students I have taught at other levels. Indeed, some of them are motivated, show interest, and work hard, some others are only prepared to do the minimum necessary, and yet some others just cannot be bothered. The only thing I like more about teaching at A-Level is that I get to know the students better, as I have far

more regular contact with them and the number of students in class tends to be smaller. On the other hand, I am greatly concerned that at both levels I cannot really 'afford' to teach the way I would like to – instead of facilitating students' own constructing of knowledge, I often fall in the trap of transmitting knowledge to students whenever I find myself hard pressed for time (and this in spite of often giving extra lessons during the holidays in order to catch up with the syllabus). Sadly, even though I consider teaching for understanding as fundamental, I do not have the courage to walk all the way indicated by Boaler (1997) in her study without the backing and support of my department.

”

6.4 Classroom Action, Interaction and Teaching Priorities

The participants tend to spend a large amount of time during the lesson next to the whiteboard, usually writing on it whilst giving their backs to the students. Andrew caricatured this reality when he said, “*When we enter the classroom, we just dash to the whiteboard and start writing on it!*” But they also stand next to the whiteboard for a very good part of the little remaining time. My classroom observation records in fact show that teachers hardly ever move away from the front of the classroom, except on the rare occasions when they move to the ‘accessible’ parts of the classroom (for there are certain parts within many classrooms that teachers cannot reach due to severe lack of space – see section 2.5.1). Usually, these ‘excursions’ either result from teachers’ desire to check that students are doing (or have done) the set work (classwork or homework) or are prompted instead by students in difficulty.

The students’ action zone during lessons is even more restricted than that of the teachers. For it is normal practice that students are seated by the time the teacher enters the classroom, and to remain seated (facing the front of the classroom) until the end of lesson. During my observations, the few students who left their seat during the lesson either did so to go out of the classroom (e.g., to go to the toilet) or to throw something into the waste paper basket. Otherwise, there is consensus amongst the teachers that, as Carmel put it, “*Ideally, students should sit down quietly and follow me*”. Both from what the teachers said during the interviews and from what I noted inside the classrooms, it seems that teachers perceive themselves and their students as having clearly defined and separate classroom roles – the teacher to lead from the front of the classroom and the students to follow from the opposite end. This physical statement positions “the authority of the teacher to be in control of the space and in

control of the mathematics” (Gates, 2002, p. 221). The teachers’ readiness to partake in the traditional practice of imparting onto students a collection of techniques that simply engage them in coming to terms with what other people have done (see Romberg & Kaput, 1999) is also evident from the type of classroom participation they value and consequently seek from their students (see Table 6.1).

Table 6.1: Student classroom participation that is valued by teachers

- **Monitoring teacher’s work:** Teachers expect students to follow closely their work on the whiteboard with an eye on identifying and pointing out any possible ‘mistakes’ (e.g., a wrong addition).
- **Monitoring other students’ work:** Teachers expect students to identify and correct the ‘obvious’ mistakes of fellow students (e.g., mistakes in a class test solution).
- **Working intermediate routine steps:** Teachers expect students to work out for them any intermediate routine step (e.g., finding the roots of a quadratic equation) whilst they are working examples for the class on the whiteboard.
- **Working examples under teacher’s guidance:** Teachers expect students to work examples in class by following closely the solution methods that they would have previously explained.
- **Asking clarification questions:** Teachers expect students to ask them questions whenever experiencing difficulty in following the lesson.
- **Answering teacher’s questions:** Teachers expect students to answer their questions, which are often ‘closed’.

The ‘passive’ participation that teachers expect from students (and teachers often complained that many students do not oblige them on this) corresponds to what I saw inside the classrooms – teachers and students operating largely in isolation from each other, hardly ever interacting in any deep or meaningful manner. One fieldwork journal entry, written when I was observing classrooms, referred specifically to this:

Although teachers and students share the same space, it seems like they live in two different worlds. My net impression is that of two sets of people regularly facing each other without actually interacting, if not only fleetingly – say, when the teacher or a student asks a question. (February, Year 3)

This situation – which shows little consideration for the understanding that students themselves construct meaning for mathematical concepts and processes, and that

classrooms consequently need to support this kind of learning (see Carpenter & Lehrer, 1999) – has parallels at the student-student classroom interaction level. During my classroom observations, I never heard a teacher ask or encourage students to work together. And whenever I saw students ‘interact mathematically’ (e.g., students comparing and discussing their work), it always seemed to result from pure student initiative, with teachers simply tolerating such behaviour. I was able to confirm this further ‘barrier’ to social interactions – which in themselves “play a fundamental role in knowledge apprehension and in skills acquisition as well as in socio-cognitive development” (César, 1998, p. 110) – inside classrooms during the follow-up interviews. Although many teachers commented that genuinely interested students stand to benefit a lot when ‘working in a group’ (often justifying this by referring to their positive collaborative experiences when still students), all of them pointed out that they do not actively encourage this practice in class. Mario crystallised the two main reasons that appear to be holding them back:

We know what students we have ... once I ask them to work together, one or two will do the work, and the rest would just copy! ... And what about the final exam? Is it fair to encourage students to work together, and then to assess them at the end of the course on something at which they have to work on their own?

That is, their first concern is what Walker and Angelo (1998) call ‘free riding’ on the efforts of the better or more industrious students by the weaker or lazier ones who, in their view, constitute the vast majority of PMI students. Secondly, they fear that should students get accustomed to such a collaborative working environment, they would not then be able to cope well with the ‘individualistic demands’ of the summative examinations organised by both school and MATSEC. The second reason – which shows that teachers are guided in their actions by the dictates of external testing (Smith, 1991) that consequently creates rather than measures what goes on inside the classroom (see Hanson, 2000) – is intimately link to what the participants identified as their top three, closely interlinked priorities in teaching. In descending order of importance, these are: (i) *helping students to pass examinations* was mentioned by all the teachers as ‘the’ top priority for reasons that had mainly to do with this being ‘the teacher’s job’, ‘what students are interested in’ (see Broadfoot, 1996), and ‘what really counts’; (ii) *finishing all the syllabus* appears to be a must for everyone, that is, except for Ray who argued that he spends as much time as necessary

on a topic, irrespective of whether or not he covers the entire syllabus; and (iii) *focusing on the better students*, namely those who stand a realistic chance of passing the examination, was mentioned by most teachers, some of whom added that albeit they feel bad about ‘neglecting’ the others, at least those who make some effort, it was simply a question of prioritising when working in a constrained environment (see section 6.3.5). Apparently, the common practice amongst teachers to base instructional decisions on only a small group of students (see Airasian, 2000; Anderson; 2003; Black, 1998) transforms itself inside PMI classrooms to include only the abler students. This focussing of energies on selecting and socialising only some students for society rather than seeking to educate all to their maximum potential reveals an understanding that “the division between those who can do mathematics and those who can’t is perfectly natural and ... legitimate” (Gates, 2002, p. 212).

6.5 The Typical PMI Lesson

6.5.1 The Teachers speak of ‘Teaching’, not ‘Lecturing’

All the participants defined their own teaching style as ‘teaching’ rather than ‘lecturing’. In particular, those with some prior teaching experience at secondary school level readily admitted that the manner in which they teach now is very similar to how they used to teach at that level. I will let Matthew explain:

I ‘teach’ not ‘lecture’ ... I don’t mention something and immediately move on to something else, like lecturers do ... Instead I go into detail, repeat as many times as necessary, and answer students’ questions ... At this level, ‘lecturing’ doesn’t pay ... students just won’t follow! They already find ‘teaching’ hard ... imagine ‘lecturing’! ... I ‘teach’ because it’s best for students. The way I teach mathematics here is not different from how it’s taught in secondary schools!

The teachers argued that whilst ‘teaching’ shows attention and consideration towards students, ‘lecturing’ lacks this human dimension. Displaying a ‘sense of practicality’ (see Hargreaves, 1994), they explained their almost complete exclusion of ‘lecturing’ (in fact, only Andrew and Kathleen reportedly ‘lecture’, albeit very occasionally) on grounds that most of the students, particularly at I-Level, cannot possibly cope with its demands. Whilst this decision is based on their awareness of what works and what does not work within their own context, it contradicts however the school’s teaching

programme that specifically includes ‘lecturing’ as an essential feature of students’ preparation for future university studies (see section 2.4.3).

6.5.2 The Two PMI Teaching Systems

The praxis within the department is for teachers to choose between either one of these two ‘teaching systems’ with PMI classes: (i) two lessons per week, and a fortnightly tutorial with half the class on a rotational basis; and (ii) three lessons per week. Only three teachers (i.e., Mario, Matthew, and Stephen) have opted for the first system. They argued that tutorials provide students and themselves the opportunity to discuss students’ difficulties on a more personal level than is normally possible during the lesson. The rest of the teachers, including myself, use the second system for reasons that have mainly to do with fears of not covering the syllabus on time.

6.5.3 The PMI Lesson – a Cycle of Exposition, Practice and Consolidation

As I was collecting the data, I became increasingly struck by the overwhelming similarities amongst the various classroom experiences in which I was immersing myself. It was as if similar stories, albeit with different emphases, were unfolding before my eyes. I wrote about this feeling in the fieldwork journal:

The more lessons I observe and the more teachers I listen to, the more I realise that in some respects they are all doing and saying pretty much the same things. This applies to how teachers tend to operate like ‘whiteboard attachments’, the very basic teaching resources they use, and even to the way they structure lessons in easily identifiable segments ... (March, Year 3)

This is not to deny that all the participants have their own particular ways of dealing with the PMI classroom situation. My point is, instead, that in spite of these individual differences, there was still so much in common in what was going on inside the various PMI classrooms to warrant my use of the term ‘the typical PMI lesson’.

The teaching resources used by the teachers included: (i) whiteboard and markers (always used by all teachers); (ii) teacher notes (most teachers had them in class, even if they did not necessarily refer to them directly during the lesson); (iii) textbook (used often by most teachers as a source of exercises); and (iv) worksheets (used only by a

few teachers, either as a textbook substitute or as an additional source of exercises). The teaching style, apart from generally adhering to the traditional ‘talk and chalk’ modality, can be normally subdivided into easily recognisable separate segments, with one phase following the other. For instance, this is how Renzo portrayed his ‘typical’ lesson:

I start the lesson by asking them about their homework difficulties. After we discuss the difficulties, I pass on to present a little bit of theory ... Then I start to work examples on the whiteboard, as many as possible ... Towards the end of the lesson, if there is any time left, I give them a question or two to work out in class ... I always end the lesson by giving them a little bit of homework.

The various phases of Renzo’s typical lesson, as well as those of the other participants, can be accommodated into a list comprising the following phases: (i) teacher presenting new theory (a feature of almost every lesson); (ii) teacher working examples related to the new theory on whiteboard (which invariably follows the new theory and accounts for the larger part of almost every lesson); (iii) students working examples in class (which occurs with some teachers only, and normally for a short period of time towards the end of the lesson); (iv) teacher setting homework (which occurs in most lessons); and (v) teacher dealing with students’ difficulties (either during the lesson, or tutorial, or both). But whilst not all the five phases in the list need occur in any one single lesson or follow necessarily this particular order, some teachers emphasise certain phases more than others. Again, it is also possible for two or more separate phases to overlap briefly at some point or other. These five phases can be further reduced to three overriding subdivisions, namely, **exposition** (i.e., teacher presents theory and solution methods, and students take down notes), **practice** (i.e., students work questions, either under the guidance of the teacher [which is mostly the case] or on their own, in order to practise the solution methods), and **consolidation** (i.e., teacher clarifies students’ difficulties). However, at an individual level, one teacher may neglect theory for methods, and another may prefer students to practise almost exclusively at home, and yet another may leave most of the consolidation work for the tutorial session. Notwithstanding this, such choices remain but individual variations on a common theme – an irrefutable cycle of exposition, practice, and consolidation that has much in common with the widely used, traditional three-segment lesson reported by Romberg and Kaput (1999). This involves

... an initial segment where the previous day's work is corrected. Next, the teacher presents new material, often working one or two new problems followed by a few students working similar problems at the chalkboard. The final segment involves students working on an assignment for the following day. (p. 4)

In essence, such lessons – which seem to ignore that, as learning is fundamentally a social process, “knowledge is located somewhere in the group space, rather than in the heads of individuals” (Jaworski, 2002, p. 73) – present ‘teaching as transmission’ and ‘learning as practice’. This reality is characterised by teachers having “to demonstrate how a manipulation is to be carried out or to explain how a concept is defined; and students ... [are] expected to memorize facts and to practice procedures until they have been mastered” (Romberg & Kaput, 1999, p. 4). Moreover, believing that their students – who they see as generally weak and unmotivated (see section 6.3.5) – offer much resistance to intellectually demanding work (see Desforges & Cockburn, 1987), the teachers disclosed a much-diffused custom in class to ‘keep calculations simple’, ‘assume nothing’, ‘go through all the steps of the algorithm’, ‘avoid technical words’, ‘do away with proofs’, ‘give little homework’, ‘leave out the harder problems’, and ‘work out most, if not each, of the classwork and homework questions on the whiteboard for the students’.

“

Michael's Story

Particularly whilst observing the lessons, I become increasingly conscious that the general PMI picture that was unfolding before my eyes largely matches my own PMI experience – I too spend much of my time at the front of the classroom, interacting directly with students only sporadically; my students' participation is similarly often 'passive'; I frequently also rely on the same set of teaching resources; and my lessons can be equally categorised into phases of 'exposition', 'practice' and 'consolidation'. It was shocking to 'look at myself in the mirror' from the back of other teachers' classrooms and realise, possibly for the first time in my life, that a good part of what I do in class “involves only the deployment of a set routine with no room for ingenuity or flair, no place for guesswork or surprise, no chance for discovery; in fact, no need for the human being” (Romberg & Kaput, 1999, p. 4). This experience has helped me to understand better that my good intentions do not necessarily guarantee good practice. The problem however remains how to break away in real classroom situations from practices of imparting knowledge to students that “are less helpful than classroom environments in which students are enculturated and apprenticed into a system of knowing, thinking and doing” (Boaler, 1997, p. 109).

”

CHAPTER 7

Teachers' Classroom Assessment Practices

7.0 The Origins of TCAP

In line with the fact that the vast majority of participants either claimed to have had no training in assessment, or else not to remember whether they had or not (see section 6.3.1), none of them – not even the three PGCE assessment ‘trained’ teachers – indicated assessment training as a source of origin for their classroom assessment practices. Ray, for instance, maintained that although he enjoys experimenting with class tests (see section 7.2.2.5 and Table 7.3), his ‘experiments’ are “*just fruit of my own mind and nothing else*”. Similar to many claims in the literature (see Cizek et al., 1995), there are instead strong indications that TCAP are largely based on the ‘experiential’ knowledge that practising teachers accumulate for themselves over the years. The evidence suggests further that TCAP are also passed on from one generation of teachers to the next. It is as if there exists a cyclic hand-over of classroom assessment practices (either from teacher to student who later becomes a teacher him or herself, or from teacher to teacher) that individual teachers subsequently modify in the light of their own classroom experiences. Stephen, who has never had any teacher training, captured the essence of this reality that is characterised by assessment knowledge being ‘caught through experience’ (see Sadler, 1989):

With regards to assessment, I practically still do as my teachers used to do when I was a student and what I saw my senior colleagues do when I started teaching. But, obviously, as I gained in experience and self-confidence, I modified these assessment practices to fit into my way of seeing things ... and, since then, I have always done the same things!

Apparently, it is the teaching profession’s communal and one’s own teaching experiences that matter primarily when it comes to devising and developing TCAP. Still, many of the participants chose to emphasise their own classroom experiences in this equation. Jackie stressed this by saying, “*I started off with my own assessment ideas, and then learned a lot more on the job!*” The fundamental role that one’s teaching experience – a characteristic that is highly valued by all the teachers (see section 6.3.2) – plays in TCAP was also evident when the teachers spoke of how they

set out initially to face the classroom assessment challenges of the then new PMI option. It was particularly revealing that every teacher who had other teaching experiences before starting to teach the PMI course (i.e., all the participants except for Kathleen and Rita) claimed that all he or she had to do to begin assessing at I-Level was to ‘transfer’ and ‘adapt’ his or her previously accrued classroom assessment practices to the new situation.

“

Michael’s Story

Although this study served to acquaint me with the world of educational assessment (see section 6.3.1), I have so far largely failed to translate my deepened theoretical understandings and resulting convictions into concrete classroom assessment practices. On the contrary, my practices remain practically rooted, like those of my colleagues, in long-accumulated personal experiences, first as a student and then as a teacher – a situation that I am finding increasingly frustrating. It is truly hard for me to accept that although I have acquired expertise in assessment, I do not even practise most of the things in which I believe and seek to promote. I blame my position – which is atypical as teachers normally have assessment policies that reflect their own values and beliefs (see Cizek et al., 1995) – on the context in which I operate. But I do wonder sometimes whether my inability to change also results from some ingrained, unacknowledged fears to leave the ‘comfort zone’ of operating within long established and collegially accepted and expected practices.

”

7.1 Roles and Relationships in TCAP

7.1.1 The Role of the Classroom Teacher

The teacher emerges from the participants’ comments as that one person who is in complete control of classroom assessment – practically deciding by him or herself about the ‘why’, the ‘what’, the ‘how’, and the ‘when’ to assess, as well as how to interpret and respond to the results of assessment. The participants’ overriding idea is that they, as professional people, have the necessary know-how to assess independently inside their classrooms. They moreover see assessment as the ‘prerogative of the teacher’ (see Brookhart, 1999; Sadler, 1989). Stephen articulated this much-diffused idea amongst the participants:

I won't let anyone intrude upon my assessment practices! I want assessment to be solely in my hands! We're all professional people ... there's no need for anyone to help us or to tell us how to assess!

Such comments made it amply clear that the participants do not particularly welcome the idea of 'others' having an assessment voice inside their classroom. At the same time, some teachers expressed deep concern that, as neither the school nor the department provides them with any guidelines regarding the computation of the Assessment Marks (see section 2.4.4), students in different classes within the same department are possibly being promoted on substantially different criteria as far as their performance throughout the year is concerned. The introduction of such guidelines thus appears to be the only 'dent' to their highly cherished and defended 'classroom assessment freedom' that these teachers are willing to accept.

“

Michael's Story

Although I do not, as is customary with many teachers, consider assessment to be strictly my personal prerogative, something that sets me apart from my students and to some extent from the rest of society (see Sadler, 1989), I admit that to date, similar to my colleagues, assessment in my own classroom is primarily a matter between the students and myself that is largely manoeuvred and controlled by me.

”

7.1.2 TCAP and 'Others'

Although all the participants spoke of classroom assessment as if it is a world that gravitates around them and their students, many conceded that other persons (most notably, the school administrators, the subject coordinator, and parents) could play an important supportive role in line with the assessment feedback given by the classroom teacher. In particular, they would like the subject coordinator and the school administrators to rebuke and punish those students who are not committed to their studies. At the same time, they expect parents to approach the teacher and to encourage their son or daughter to work harder, especially when a low Assessment Mark warrants this.

“

Michael's Story

To designate classroom outsiders (be they colleagues, administrators, or parents), as many colleagues have suggested, to act simply either as a deterrent or supportive force would only serve, in my view, to reinforce even further the present isolationist framework that characterises our assessment practices. My efforts within the department to create a more inclusive assessment environment (see 'My Journey as a Teacher Continues ...') attest to my belief that only a concerted effort by all concerned can help transform our closed assessment culture into a more open and beneficial one (see NCTM, 1995).

”

7.1.2.1 TCAP and Colleagues

All the participants, except for Rita, claimed that in matters related to classroom assessment they only sought the advice of other teachers (usually, the more experienced ones) or collaborated with other teachers (usually, the other beginning ones) when they were still in the initial stages of their teaching careers (see Bezzina et al., 2004). But, as Matthew pointed out, this 'assessment collaboration' between teachers does not normally continue beyond the settling down period (which, according to the participants, may take up to one year):

It is normal for new teachers to work together and to approach the more experienced ones to help them with assessment ... I did this when I started off! But, really, I see no need for teachers to work together on assessment once they would have settled down ... once they are no longer green!

By the time of the fieldwork phase (see section 5.3), all the participants had already been at the department for at least two years. This 'having already settled down' probably explains why I found very little evidence of any assessment collaboration or even exchange of ideas amongst the teachers (say, the sharing of the Assessment Marks criteria). Angelo captured the essence of this highly individualistic classroom assessment reality (see Cizek et al., 1995; Hall et al., 1997; Stiggins & Bridgeford, 1985) when he said:

I don't actually know how the other teachers assess in their class ... I can only guess! After all, it's not something that concerns me, is it? I don't know about them, nor do I tell them about myself ...

Rita, by virtue of her ongoing assessment collaboration with non-participant Francesca (which extends to other professional areas as well – see section 6.3.4), offered the only notable exception to this non-collaborative environment. I will let her explain:

I work a lot with Francesca ... ever since we did the BSc together! Our collaboration doesn't concern the PMI option, which she doesn't teach ... But at A-Level we do everything together ... we plan and give the same class tests and worksheets, we discuss and share the examples we use in class ... We don't work together throughout. We first work on our own, and then we meet to discuss the material we selected individually ... then we decide what to keep or leave out. Although I feel comfortable working with Francesca, there're still things on which we don't agree ... when this happens, we just go our separate ways ...

Their assessment collaboration thus appears to be mainly the product of two persons who had grown accustomed to working together prior to setting foot within the department. And whilst they both reportedly value and cherish their teamwork, each of them – just like the other participants – has retained full control on her own classroom assessment practices (see also section 6.3.4). But then, even though none of the other teachers actually collaborate with each other on classroom assessment, there is evidence to show that, albeit only occasionally, some do approach colleagues from their office on assessment related matters. These ‘flashes’ of assessment interactions may be grouped under the two headings identified in Table 7.1. Notwithstanding these ‘flashes’, the general lack of teacher assessment collaboration was also evident when some participants primarily attributed their knowledge of colleagues’ assessment practices (however limited) to information passed on to them by students.

Table 7.1: Forms of limited assessment interaction amongst teachers

- **Sharing of information:** This involves the exchange of student assessment records between first year and second year teachers at the start of the scholastic year, and any subsequent informal discussion amongst teachers, say over a cup of coffee, regarding the academic progress of former or current students.
- **‘Discussing’ class tests and their results:** This involves showing the class test (usually when already finalised and ready for photocopying) to some colleagues. Only very rarely are there any changes, significant or not, in the test paper following this ‘consultation’. After marking the test scripts, the teacher informs colleagues about the ‘quality’ of his or her students by publicising the student marks (at this point, the test paper is usually shown again to colleagues so that they may better interpret the test marks).

In this non-collaborative environment, trust amongst teachers appears also to suffer (see Cole, 1997). This emerged when some participants made certain ‘insinuations’ about a number of colleagues. These included: (i) giving easy class tests that consequently inflate the Assessment Marks (see section 7.3.3); (ii) not knowing how to set a ‘proper’ class test (or the end-of-first-year examination) paper; (iii) resorting to outright favouritisms when giving the Assessment Marks; and (iv) fiddling with the Assessment Marks of the weaker students in order to ‘protect’ their reputation. The last point highlights in turn the rather diffused fear amongst the participants of being judged negatively by colleagues. This is reportedly why, for instance, Mario does not give any ‘easy’ test items (as this may give the impression that he sets low standards in class) and Carmel is reluctant to publicise his test and examination results (as he is afraid that colleagues would think badly of him because of them).

“

Michael’s Story

My closest working relationship at school is with Jackie. We joined school together and have been in the same office ever since. But it was probably our common inexperience at sixth form level that actually drove us to ‘interact professionally’. In the first year or two, although we actually planned and worked on our own, we still used to frequently consult each other on matters that had mostly to do with the interpretation of the syllabi and the clarification of the finer details of the mathematical content. However, our interactions, which continue to the present day albeit in a less intense form, have had (and still have) little to do with classroom assessment. In this respect, apart from the ‘small talk’ that occurs spontaneously amongst friends and colleagues, we continue at most to review each other’s tests before giving them to our respective classes. Not even with Jackie have I ever specifically discussed the criteria I use to compute the Assessment Marks. And even though she thought otherwise, I actually got to learn about hers from the ‘assessment interview’. My assessment interactions with the other colleagues are at an even lower level. In fact, were it not for this research, I would have most probably remained largely ignorant of the participants’ classroom assessment practices (as they must still be with regards to mine).

”

7.1.2.2 TCAP and Students

Although the participants see students as the only persons who, apart from themselves, are directly involved in classroom assessment, they still relegate them to a clearly subjugated role, and retain a firm hold on all significant classroom assessment decisions in their own hands (see Cross, 1998). Students are thus on the receiving end

of assessment, simply expected to follow the assessment lead indicated by their teachers. Suffice it to mention that the only recorded instances of students being directly involved by their teachers in classroom assessment decisions concerned the selection of class test dates. Jackie provided a glimpse at this reigning mentality:

I'm in charge of classroom assessment ... it is I who decide when and how to assess ... I expect my students to cooperate with me. I expect them to answer my questions in class, to tell me when they have difficulties, to do the work I set them ... and to study when they have a class test coming up ...

Students' self-initiated participation in classroom assessment matters is largely restricted to when they make their difficulties known to the teacher or else make pertinent observations or comments (see section 7.2.2.3). Notwithstanding the participants' overwhelming desire for students to 'reveal themselves to them' in this manner, many teachers – who appeared to be particularly into the 'coverage game' (see Ellis, 2001) – also maintained that they are ever so careful not to let these interventions seriously affect the flow of the lesson, as this could compromise the syllabus coverage. Andrew was one of the teachers who spoke in favour of striking a balance between student interventions and sticking to the pre-lesson plans:

I'm all in favour of giving students time to ask questions ... I cannot help them unless I know what their difficulties are! But there's a limit to everything ... by the end of the lesson I would want to cover most, if not all, of what was in my plan ... My students have come to understand this ... they know that at some point I may have to tell them, 'Sorry, but I cannot take any more questions!'

To this effect, I noted a common strategy amongst the participants to largely limit student interventions to times when they are least likely to falter the flow of the lesson – that is, basically to periods after the teacher has finished a self-contained part or all of the 'exposition' (see section 6.5.3). The unilateral nature of TCAP is also reflected in the manner in which the teachers relate to students with regards to the formulation and divulgation of their Assessment Marks criteria (see section 7.3.3). For whilst all the teachers reported that they neither negotiate nor consult with students in formulating these criteria, only half of them claimed to indicate clearly to students what these criteria are. The rest reportedly provide either vague indications (two teachers) or no indications at all (four teachers).

TCAP also relate to students in the following two manners: First, the application of practically the same practices with all students in class. And second, the intention for practices to target students primarily as individuals. Starting with the ‘sameness’ of TCAP, my probing into whether or not the participants use the same assessment practices with all their students in class met with many startled looks. Except for Ray, the possibility of assessing different students differently emerged as an unacceptable novelty to all participants. But even though Ray argued formatively that as all students are individuals, one should ideally assess them in a manner that best suits their individual needs (see Black, 1998), he still admitted not to practise differentiated assessment due to time constraints and fears of appearing unfair to both students and classroom outsiders. The other teachers – apart from acting summatively like Ray by also linking classroom assessment to grading and wanting to produce publicly defensible grades at all costs – also judged the eventuality of ‘non-sameness’ in assessment as intrinsically unfair. This is how Andrew captured these concerns:

Besides not even knowing how this can be done ... would it be fair? ... Imagine grading some students on class tests and others, say, on coursework ... If I were a student, I would surely not approve of my teacher doing such things! I think it is only fair that all students are graded on the same forms of assessment.

On the other hand, the participants’ declared preference for ‘having each student working on his or her own’ (see section 6.4) is reflected in the manner in which they go about gathering assessment information (see section 7.2.2). For they act throughout in a manner that does not foresee the need for (or encourages) any interaction amongst students, and repeatedly target students only as individuals.

“

Michael’s Story

Like my colleagues, I practically relegate my students to the receiving end of assessment. Their self-initiated participation is in fact limited to: (i) being able to consult me at my office whenever the need arises (see section 7.2.2.2); and (ii) almost having a free hand to intervene with comments and queries during the lessons (see section 7.2.2.3). But then, albeit I consult them on test dates and indicate clearly the standards I expect from them and the Assessment Marks criteria I use, I do not negotiate either the ‘standards’ or the ‘criteria’ with them. Again, although I try to be as sensitive as possible in my assessment practices to the individual behavioural characteristics of students (e.g., questioning a student only on a one-to-one basis once I realise that he or she ‘hates’ being questioned in public), I do not actually organise my practices according to their individual assessment

needs. Apart from lacking the time needed to venture properly in this direction, I would probably also find it hard to abandon my limited 'set menu' of practices that I developed over the years, and with which I feel comfortable still. Finally, most of the assessment information that reaches me in class, similar to that of my colleagues, originates from and reflects the efforts of individual students. The only notable exception to this is when I occasionally present the students with an open situation, give them a couple of minutes to discuss it amongst themselves in small groups, and then engage in a wider 'discussion' with them (see section 7.2.2.4).

”

7.2 Ways in which Teachers obtain Information

The participants distinguished between two separate sources of information from which they draw insights and understandings into what goes on inside their classroom. The evidence obtained can be external to both themselves and their classroom, or it can stem directly from inside their own teaching-learning situation. The teachers thus have two ways of obtaining information. First, they can collate any pre-existing information (something they normally do at the beginning of the year). Second, they can generate their own information (which is an ongoing process throughout the year).

7.2.1 Collating Information

Some participants complained that they do not even have access to the readily available, albeit scarce, school student records that are linked to the PMI option (i.e., the O-Level mathematics grades, the first year Assessment Marks, and the end-of-first-year examination results), let alone to other student records that predate the students' arrival at school. As a matter of fact, the only information about students that a teacher is provided with at the beginning of the scholastic year is a class list (with just the student names on it) that is designed to record student attendance. To get any additional information, the teacher will either have to ask the students themselves, or else to approach, should this apply, the colleague who was in charge of his or her class during the first year. The information obtained from these requests, which see teachers putting up their 'antennae' at the start of the year in search for information (see Airasian, 2000), includes: (i) *personal details*, such as, home address and telephone number, school last attended, and the other options of the Matriculation Certificate; (ii) *assessment records*, such as, O-Level mathematics grades, the first year class test

marks, homework marks (in the few instances when they exist – see section 7.2.2.2) and Assessment Marks, and the end-of-first-year examination marks; and (iii) *oral accounts*, such as, when a colleague provides anecdotal information about a student.

In reality, though, only a few of the participants obtain such records and oral accounts. Although some mentioned ‘feeling uncomfortable’ with having to ‘beg’ students and colleagues alike for this information, the main reason appears to be their general refusal to use or trust externally sourced information (see Cizek et al., 1995; Hall et al., 1997). The teachers claimed to treat such evidence, when and should it be available, either as a ‘vague indication’ of what to expect during the year – something with which to ‘wet their appetite’ before they see things for themselves – or else as a point of comparison and/or contrast with what their evidence is presently telling them. In a scenario characterised by diffidence towards externally sourced information, teachers are starting at best each new scholastic year with only the barest of assessment information about their students. Moreover, those of them who pay no attention to the assessment records of previous teachers (see Weeden et al., 2002) are possibly facing students without even knowing anything about their assessment background.

“

Michael’s Story

As I teach second year PMI classes, in the beginning of the year I try to muster whatever information I can – including any available assessment records – about my ‘new’ students from their ‘old’ teacher. I also ask my students to fill in a data sheet that includes personal details, the O-Level grade and the end-of-first-year examination result. Like my colleagues, however, I do not attach too much importance to this type of information. But this has nothing to do with not trusting external sources. It has to do instead with my understanding that not only do students change over time, but that they also react differently to different learning environments. I am thus primarily interested in external information in as far as it can help me to gain a better understanding of my own information.

”

7.2.2 Generating Information

All the participants claimed to be by far more convinced of and ready to rely on any evidence that is generated directly from within their own classroom than anything else

that originates from outside it (see ‘tried and tested practitioners’ of Gipps et al.’s [1995] intuitives model). Apart from believing to be as teachers in a position to gather unique information about students (see Calfee & Masuda, 1997), the participants invariably linked their strong faith and reliance on classroom-based information to what they view as their own sustained active involvement with students at close quarters. This came out strongly in their constant reminders that they ‘know things’ because they ‘teach’ rather than ‘lecture’ (see section 6.5.1). The teachers evidently believe to be practising a teaching style that intrinsically guarantees them knowledge. Nicholas clearly elucidated this point:

I can speak confidently about my classroom situation because I’m constantly overseeing what’s going on ... I’m all the time in close contact with students. As I’ve already explained, I teach not lecture ... If I were to give a lecture, I wouldn’t even know a small fraction of what I know about my teaching and students!

The ways in which the participants claimed to generate information about their respective teaching-learning situations can be classified under five categories: (i) observing students; (ii) checking students’ work; (iii) listening to students; (iv) asking questions; and (v) class tests.

“

Michael’s Story

Like my colleagues, my attention is mostly focussed on the information generated within my own teaching-learning situation. But whilst I consider my choice to ‘teach’ rather than to ‘lecture’ as contributing towards the richness of the accessed information, I remain aware that what I come to know is characterised by both partiality and inexactness (see Broadfoot, 1996; Harlen, 1994b; Nuttall, 1989). Thus, albeit each infinitesimal piece of information helps me to gain a better understanding of the evolving picture, there is no way that my knowledge – which I also generate in the manner described by the other participants – is ever going to be complete, or even close.

”

7.2.2.1 Observing Students

Most of the teachers spoke about the importance they attach to observing students (see Airasian, 2000; Stiggins & Bridgeford, 1985). They explained that by keeping a vigilant eye on the students, the teacher would be able to get insights into the

effectiveness of his or her teaching, that is to say into whether or not the students are learning. Many of them maintained further that, very often, it is enough for them to look straight into the students' faces to get an indication of their level of understanding. It was Nicholas who elaborated mostly on this aspect:

I'm able to tell if a student understands from his face! If he tries to escape my glance, it's a sure sign that he hasn't understood ... If he stares back at me, he's probably lost ... If he looks at me, then lowers his head, and then looks at me again, he's still not convinced ... I cannot say that I'm always correct about these things, but usually I am!

Nicholas, like most of the participants, appeared somewhat aware that any conclusions reached from observations could be misleading (see Anderson, 2003). In particular, given the teachers' propensity to remain close to the front of the classroom throughout the lesson (see section 6.4), they are very likely to be observing most of their students from a considerable distance. But some manage to narrow this distance when they occasionally move around the classroom, at least up to those parts that they could reach in the generally overcrowded classroom environments (see section 6.4). In either case, the teachers commented that they learn a lot when they observe students at work (usually working on their own – see section 6.4), at least those they can reach. I will let Kathleen explain:

I learn a lot when I go round the classroom and see my students working ... I get to know how they work, where they are having problems, if they are following what we're doing in class ... I just wish I could go more often round the class ...

Most of the participants, however, reported that they tend to focus their 'spontaneous observations' (see Airasian, 2000) in class, both when they teach and when the students are at work, primarily on the better students (see section 6.4). These high achieving students reportedly act as the 'steering group' that helps the teacher to set the pace of the lesson – moving on when they appear to be following, and holding back when they appear to falter (see Anderson, 2003). Jackie was one of the teachers who spoke more openly about this reality:

I pay most attention to the better students in class ... I'm constantly on the look out ... looking at their faces ... to make sure that they are following me in class. Even when I go round the class to see students working, I'm mostly concerned to find out what they're doing ... if they're with me, I know that I can move on.

These participants also concurred that sometimes it is best for the teacher to turn a blind eye to his or her observations of the unmotivated students, as the teacher would gain on time (which they see as more rewardingly spent on the better and motivated students) by avoiding sure conflict with them.

“

Michael's Story

Although I gain through observation an inkling of what is going on inside the classroom, it does not actually reveal anything about what goes on inside my students' heads. Still, whenever space and time permits, I like to move about the classroom in order to observe closely students at work – the type of observation that I find most rewarding. Irrespective of their ability or performance, I try to keep a vigilant eye on all the students. However, I find it very hard to 'really bother' with those students who make it very obvious that they are only in class in order not to lose out on their maintenance grant (see section 2.1.1).

”

7.2.2.2 Checking Students' Work

Some participants included 'checking students' work' amongst their classroom assessment practices. However, it appears to be the least practised (both in terms of the number of practitioners and frequency) and valued assessment tool through which they learn about their teaching-learning situation. The few teachers who reportedly indulge in this practice are engaged at two different levels. At level one, the teachers simply try to verify if the students have done, or at least attempted, the work they set them. They normally do this during the lesson itself either by selecting and checking the work of some students, or by quickly going around the classroom to check everyone's work, that is, time and space permitting. Jackie, one of the few teachers involved in this practice, explained the assessment benefits of doing these 'spot checks':

When I glance at students' work, I get immediate feedback about my teaching and their learning ... If most of them – at least those whom I'm expecting to follow me – have done their work, then I can confidently move on to the next step ... but if I notice serious difficulties, my priority then would be to consolidate what we have already done.

Many of the teachers, however, judged this practice to be a waste of time. Their reasons had mainly to do with the much-diffused opinion amongst them that PMI students are mostly untrustworthy (see section 6.3.5). This is what Carmel said:

Students are likely to copy the homework once the teacher insists on seeing it ... that's why I never bother to check their work in class! I would be wasting my time, and so would they. I'll rather have students doing only a little bit of homework, instead of copying all of it for my sake!

Even less teachers than those involved in 'level one' checks claimed to have either the will or the time for 'level two' checks. At level two, teachers collect a selection of students' work (say, a worksheet or a textbook exercise) to go over it in some detail either at home or at the office. Notwithstanding their general reluctance to carry out such checks, many participants conceded that, given the 'right circumstances', such a marking exercise would provide them with important insights into their teaching and students' learning (see OFSTED, 1998). All the teachers who opt out of marking explained this decision by claiming either that most of students' work is copied, or that there are too many students in class to make this a feasible assessment reality, or both. In truth, only Ray, Rita (who was seriously thinking of quitting what she is increasingly seeing as an 'unyielding' practice) and Kathleen reported to collect and mark their students' work. This occasional exercise (my document analysis suggests between three and five times a year) invariably consists of a worksheet normally given to students over the weekend or holidays in order to minimise as much as possible the interaction amongst them. In a reality that more than echoes OFSTED's (1998) preoccupation about the low frequency and quality of marking, the rest of the teachers, unless specifically approached by students with classwork or homework difficulties, only come into direct contact with their students' work when correcting class tests.

“

Michael's Story

I never specifically ask students to show me their work – considering their age, I find this approach a little childish. Instead, I try to instil in them from our first meeting a sense of personal responsibility. I point out that I do not check on their work, not because I cannot be bothered, but because I prefer them, as young adults, to start taking some direct responsibility for their own studies. In the first year or so, I used to occasionally collect (I always notified the students about the collection dates), mark and grade a selection of their work (usually a worksheet). But I stopped doing this as I learned from experience that most

of them, instead of working together outside the classroom as I had hoped and encouraged them to do (but I did ask them to indicate the persons with whom they worked), preferred to copy the set work and pass it off for their own. Nowadays, I still manage to get a close look at students' work whenever they approach me individually or in small groups during personal contact hours (see section 2.4.1). But whilst some students make good and constant use of my availability – which goes beyond the official times established by the school – some others never bother to. At times I wonder if some students are mature enough to handle the responsibility that my approach demands.

”

7.2.2.3 Listening to Students

A much-quoted source of information amongst the participants was ‘listening to what the students have to say’. There appeared to be a general consensus amongst them that by listening to students, they can ‘get inside’ their students’ heads and thus obtain useful insights into their teaching-learning situation. Other than when students are asked questions (see section 7.2.2.4), teachers claimed to ‘listen’ when students either raise difficulties or else make pertinent observations or comments. This ‘listening’ occurs during lessons (which was by far the most frequently cited period), tutorials (which only three teachers use – see section 6.5.2) and personal contact hours (which, according to most teachers, are only very rarely used by PMI students). In particular, ‘listening’ seems to occur during all the three phases of the PMI lesson (see section 6.5.3). This is how Andrew related ‘listening to students’ to the ‘consolidation’, ‘exposition’ and ‘practice’ phases of his lessons:

At the start of the lesson I invite them to raise their difficulties ... these may concern the previous lesson or the homework [consolidation]. Although I don't particularly like to be stopped when I'm delivering the lesson – I'd rather they'll ask me at the end of my delivery – I still allow them to stop me and ask for clarifications when they have serious problems [exposition]. ... When I'm working examples or when they're doing practice work in class, I expect them to confront me with their difficulties ... I don't want them to leave the classroom with lingering doubts [practice].

In line with their preference for ‘teaching’ (see section 6.5.1), all the participants allow the students (some even spoke of ‘encouraging’) to ask for clarifications, pass comments, and raise difficulties during lessons. The teachers, in turn, not only expect students to ‘reveal themselves’ to them through these interventions – which reportedly do not originate either from discussion or collaboration amongst students, especially

inside the classroom – but are also of the opinion that students’ failure to do so (which, according to the participants, is often the case) seriously jeopardises teachers’ efforts to address properly the underlying problems.

The teachers invariably linked ‘listening’ to receiving feedback from students within reasonable time, often there and then. But in spite of the ‘immediacy’ offered by listening (see Weeden et al, 2002), many teachers confided that they are far from enthusiastic with the actual results they reap from listening to students. Their complaints were largely related to the fact that only the same few students raise difficulties in class, and that hardly any student appears interested enough to enter into a discussion with them beyond the bare necessities (e.g., finding the correct solution method for a particular question). Some teachers claimed to find this reality – characterised by only very few words of classroom talk being actually spoken by students (see Torrance & Pryor, 1998) – most frustrating. Matthew is a case in point:

I’m not saying that most of my students do not pay attention in class, or that they do not copy from the whiteboard ... nothing like that! Most are well behaved in that respect ... but the problem is that this is all they do! They’re so passive ... most of them don’t even bother to ask me their difficulties! It’s so irritating!

And just as the participants tend to observe more attentively their better students (see section 7.2.2.1), so do most of them give more importance to ‘listening’ and ‘reacting’ to what these students have to say. Jackie is again one of these teachers:

My better students know from experience that they can stop me whenever they have a query, and that I would listen to them and take all the necessary time to clear it out. But I don’t do this with the weaker ones ... there just isn’t enough time! When a weak student asks me something, I answer hurriedly, and move on! ... I think the weaker ones realise that I behave differently with them ...

“

Michael’s Story

‘Listening to students’ not only gives me the opportunity to learn a lot about my teaching-learning situation, but also to do so in the course of the lesson itself – something that permits me, in theory at least, to immediately adjust or redirect my teaching efforts accordingly. Sadly, though, most of my students choose not to intervene at all in this manner during the lesson. And this in spite of my disposition and continuous reminders that they can ask questions or pass comments whenever they feel the need, as long as this is done in an orderly manner. When circumstances permit (e.g., a question or comment lends

itself to this), I seek to turn the students' already rare interventions into a wider class discussion. But such discussions do not often ensue, as their interventions are usually of a very closed nature. Hard as I try to give each student the same level of attention, I too have my own 'invisible' students with whom I have very few direct interactions and about whom I have much less personal knowledge (see Weeden et al., 2002).

”

7.2.2.4 Asking Questions

Many participants commented that ‘asking questions’ and ‘listening to students’ very often occur during the same interaction between the students and themselves, with one often leading to the other. According to some of them, the only real difference between the two is that ‘asking questions’ inverts the initial active agent from student to teacher. All however concurred that ‘asking questions’ offers the teacher the same assessment opportunity as ‘listening’ – that is, the possibility to learn about his or her teaching-learning situation from what the students have to say. The underlying strong belief amongst the participants is that when students answer correctly to teacher’s questions, the teacher has indications that he or she is teaching effectively as the students appear to be learning. Ray exposed this line of reasoning:

I'm all the time asking questions in class ... when I ask and they answer back correctly, I would know that they have understood ... I would know that my teaching has been successful.

Their declared belief in the potential of teacher questioning to reveal if students know or do not know something – which probably explains why they all hold ‘asking questions’ as one of their more important assessment tools – contrasts with the understanding that questioning cannot guarantee to elicit evidence of knowledge (see Brookhart, 1999; Pryor & Torrance, 2000; Wiliam & Black, 1996). Whilst everyone, except for Carmel and Kathleen, spoke confidently of his or her ability to ask questions, many complained that were it not for the usual few interested students, they would hardly be able to draw any feedback at all from their questions. They also claimed that most students do not even make the effort to ‘think’ when answering questions. But then – in line with the expressed concerns about the quality of teacher questioning (see Black & Wiliam, 1998; Stiggins et al., 1989) – both the interviews

and my classroom observations reveal that teachers' questions are predominantly closed. This is, for instant, how Nicholas described one of his typical questions:

I'll tell them, "Think about this problem! If the function is positive for all values of x , what can you tell me about its roots? Think about it!"

Apart from simply evoking memory rather than challenging students to think, questions put by teachers came out as more likely to be addressed to the whole class (with students, in what is clearly the result of individual minds at work, answering in chorus) than to individual students. Moreover, teachers emerged as very likely to answer their own questions when students fail to answer almost immediately. The following classroom observation excerpt (showing Stephen whilst working Coordinate Geometry practice questions with students) indicates this emphasis on speed, as opposed to thinking deeply about things, culminating in students being passed over by the teacher (see Boaler, 1997; Lesh et al., 1992):

Stephen: *What can you tell me about the gradients of two perpendicular lines?*

Students: (some, in chorus) *Their product is equal to minus one!*

Stephen: *That's right! Now, how do we find the point of intersection of two straight lines?*

Students: (no immediate reply) ...

Stephen: *By solving simultaneously! Have you already forgotten?*

A good number of teachers disclosed that they avoid questioning directly those students who are unlikely either to answer correctly or else immediately. There appears to be two main reasons for this. First, and this was the commonest explanation offered, there are deep fears linked to 'exigencies of teaching' that the direct 'inclusion' of these students would inevitably disrupt the lesson flow (see Broadfoot, 1996; Torrance & Pryor, 1998). A distant second was the possibility that such students would face public embarrassment as a result. A compromise position between outright 'inclusion' and 'exclusion' also exists. As Nicholas explained, this involves gauging the level of the question according to the perceived level of the student concerned:

If I know that a student is weak, I'll just ask him an easy question ... something he is likely to know! I leave the more challenging questions to the better students ...

But, irrespective of whether the teachers' questions can be classified as 'easy' or 'difficult', there is hardly any evidence to suggest that they can and do lead to the

stimulation of original thinking in students. The teachers appear to be well aware of this. Indeed, many commented that their habitual participation in the traditional ritual sequence of ‘question by teacher, response by student(s), and feedback/evaluation by teacher’ is not a discussion in its true sense (see Torrance & Pryor, 1998).

“

Michael’s Story

Although ‘asking questions’, just like ‘listening to students’, provides me with instant feedback on my classroom situation, I am highly critical of the feedback quality I receive. But given that my forays into ‘genuine questions’ (see Torrance & Pryor, 1998) remain occasional, I cannot honestly expect any better. Moreover, even though I repeatedly tell myself that as student interventions in class are not very forthcoming (see section 7.2.2.3) I ought to make up for this through my own questions, I still often end up sacrificing ‘question time’ to favour the ‘exposition’ and ‘practice’ phases of the lessons. Again, the closest I come to engaging students in a real discussion is when I occasionally present them with an open situation, which either arises from the lesson itself or which I would have prepared beforehand, and give them a couple of minutes to discuss it amongst themselves in small groups in an effort to map out a solution. This is followed by a brief presentation and analysis of the various ‘solution maps’.

”

7.2.2.5 Class Tests

The one assessment practice to which all the participants kept repeatedly referring to when we discussed how they learn about their teaching-learning situation was ‘classroom testing’. It is as if they cannot imagine classroom assessment without tests. The teachers explained their great attention to testing by arguing, from what is clearly a behaviourist perspective, that tests can provide a ‘very good picture’ of what individual students know at a particular time (see Torrance & Pryor, 1998). Albeit they all emphasised the ‘need’ to test students regularly in order to better shed light on their teaching-learning situation, Table 7.2 (which shows the number of class tests that the participants gave during Year 2 of the study) suggests that different teachers might have different notions of what ‘regular testing’ actually means. For whilst the extra first year term (see section 2.4.3) might account for the differences in the number of class tests between the first year (an average of 5.7 tests per year) and second year classes (an average of 2.8 tests per year), the differences between same year classes cannot likewise be so easily explained. With regards to the content of the tests – which

normally feature only one mathematics topic (e.g., a class test on Integration) – the teachers claimed that quite often they choose tests items that are very similar (sometimes almost word for word) to the questions that they would have previously

Table 7.2: Number of class tests during Year 2 of the study

1st Year Class	Andrew	Angelo	Jackie	Mario	Matthew	Nicholas	Stephen
Class Tests	9	6	4	6	5	5	5
2nd Year Class	Carmel	Kathleen	Ray	Renzo	Rita		
Class Tests	2	2	5	3	2		

presented and tackled in class (see section 7.4.3 for a discussion on the nature of these classroom questions). The practice of essentially pre-laying out the reasoning involved in solving class test questions means that most of these, even those that may appear to require complex thinking, simply engage students in ‘recall experiences’ of things previously done in class (see Brookhart, 1999). Ray recounted the most extreme case of this testing characteristic encountered in the study:

A couple of weeks before a class test I give them a worksheet for homework. First, they work it out, and then they can ask their difficulties in class ... who’s interested enough can even take note of how to work those questions he has difficulty with! Then, for the class test I set a paper that is very similar to the worksheet ... it would have the same type of questions! For example, if in the worksheet they had to find the number of permutations from a five-letter word, in the test they would have to find the number of permutations of a different five- or four-letter word ... the tests questions would require the same thinking!

The paper-and-pencil class tests, which normally take the full lesson hour, are held under strict examination conditions – a way to ensure, the teachers repeatedly stressed, that by the end of the test they would be able to determine what each individual student can ‘produce’ on his or her own. Many teachers argued that this ‘characteristic’ – which supposedly excludes ‘free-riding’ (see Walker & Angelo, 1998) – enables them to feel surer of what they know about students from a class test than from any other form of assessment. This ‘confidence’ parallels their trust in their ability to ‘measure’ how students perform in a test and to present this under the guise of a mark, usually as a percentage. This behaviouristic emphasis on measurement was,

albeit to a lesser extent, also on Ray's mind – the only teacher to depart occasionally and partially from the 'traditional' testing arrangements of the other participants. One of his class test 'experiments' evidences this (see Table 7.3).

Table 7.3: A class test 'experiment'

During one of my classroom visits, Ray informed the students that they could prepare handwritten notes on one size A4 sheet of paper that they would then be able to bring in with them and refer to during their forthcoming class test. Although it was my intention to quiz him at some point in the follow-up interview about this practice, he preceded me and insisted that he should clarify what it was all about. He explained the rationale behind his 'experiment' as follows:

It's a new thing I'm trying ... My idea is that if they have to write notes on just one sheet of paper, they'll have to choose what's really important, and to do this they'll have to go all over their notes! I think this encourages them to do some study for the test because they know that it's likely to pay off. But they'll have to be very selective ... they'll have to cram on the sheet only what they see as very important, and leave out what they judge as less important or what they already know ... This will also help them for the exam preparation, as they'll be learning how to write short study notes. But I've already told them that for the next test I won't be letting them bring in any notes ... it's just a one off! If I'll let them use notes in every test, they won't really know where they stand in their exam preparation ... most students agreed with me on this! Two of them even said that they wouldn't be using notes ... not even on this one occasion as the mark they'll get won't be 'truthful' ... and they're right!

“

Michael's Story

Although the understanding that classroom assessment should be essentially about the promotion of learning convinces me that class tests ought no longer enjoy the 'high profile' in schooling that they have (see Ellis, 2001), I still normally give, like most of my colleagues who teach second year PMI classes (see Table 7.2), two or three class tests each year. My tests, which largely follow the same parameters outlined above, remain rather traditional fare. Lately, however, I have started to wander away (just like Ray!) from the strictly conventional testing formula that is so domineering within the department. My little excursions have included 'open book tests', 'Ray's experimentation with notes', and "helping' students during the test'. However, some students (just like Ray's) have complained that such 'allowances' do not let them assess properly where they stand in a 'real examination'. I have since given students the option to opt out, should they so desire, from using similar 'allowances'.

”

7.3 Formal and Informal Sources of Information

In the course of the interviews, a distinction invariably emerged between sources of information that lead directly to student marks and others that do not. I decided to refer to TCAP that produce student marks as ‘formal’, and the rest as ‘informal’. Applying this terminology, the participants’ only two formal sources of information are ‘class tests’ (which applies to all teachers – see section 7.2.2.5) and ‘checking students’ work’ at level two (which only applies to three teachers – see section 7.2.2.2).

7.3.1 Information gained from Formal and Informal Sources

Whenever the participants spoke about their informal assessment practices, a sense of the ‘there and then’ consistently came to the fore. In fact, the teachers not only described these practices as ‘spontaneous’, ‘immediate’ and ‘unofficial’, but also concurred that they are largely dependent on ‘teacher instinct’. On the other hand, the teachers presented their formal assessment practices as planned official activities that require subsequent ‘intervention’ on their part (read ‘marking test scripts and worksheets’) before any information becomes available (see Stiggins & Bridgeford, 1985). By and large, the teachers maintained that, albeit in different ways, both the formal and informal assessment practices are useful classroom tools. With regards to informal assessment practices, all the teachers – except for Carmel who reportedly finds it hard to establish ‘contact’ with students during lessons – claimed that they provide them with a comprehensive real time picture of their general teaching-learning situation that contributes significantly towards the smooth running of the lesson. But whilst this picture helps teachers to form and to improve ongoing classroom processes (Airasian, 2000), they argued that for a more individualised picture of students they need to assess students informally over an extended period of time – a sort of gradual accumulation of knowledge (see Watson, 2001). Nicholas spoke at length about this:

The feedback I get during the lesson is fundamental to my teaching ... by observing the students, listening to their difficulties, and ... even by looking at their expressions ... I get an overall idea of how my lesson is being received ... whether they, at least most of them, are following me or not. There is no way of knowing there and then what each individual student is doing or thinking ... All I get is a general impression which helps me to move forward my lesson ... helps me decide if I should repeat something, give more examples, or move on to

something new ... things like that! I only get to know students individually over time ... and I still get to know some more than others ...

The teachers' day-to-day dependence on informal assessment (see Broadfoot, 1996) emerged clearly when all of them, again except for Carmel, indicated to find their informal practices to be more useful in everyday teaching than the formal ones.

Whenever the teachers referred to the usefulness of their formal assessment practices – which reportedly have less immediate application than the informal ones – everyone focussed almost exclusively on class tests. This suggests that, even though a few teachers also obtain marks from corrected worksheets, they all overwhelmingly consider class tests as ‘the’ formal source of assessment information. The teachers concurred that the concise and comprehensible test marks, a testing by-product that they value highly, normally provide a fairly good idea of the students’ general understanding and knowledge (see Brown, 1990). This understanding – apart from leading many participants to assert that their gradual accumulation of tests marks provides them with a clear indication of students’ likely examination performance (the ‘veracity’ of these claims is examined in section 7.6.1) – possibly explains their strong commitment to continue using class tests, and why some of them use test results to chart students’ progress over the year.

“

Michael’s Story

My informal practices are far more useful to my everyday teaching than the formal ones (i.e., class tests). For it is on the basis of their ongoing information that I navigate through each single lesson, and plan from one lesson to the next. Informal assessment provides me with immediate feedback that, albeit of a very general nature, permits me a measure of flexibility. The main drawback, however, is that information from informal sources is hardly ever focused or representative enough – more than anything else, it provides a series of often unrelated snapshots of the classroom assessment reality. In truth, I can only draw vague indications of the overall teaching-learning situation in the course of the lesson. On the other hand, class tests, even if few and far in between, provide me with a deeper understanding of particular students. Albeit there are serious limitations to what tests can actually reveal (see Gipps & Murphy, 1994), I find that properly designed class tests give rather good indications of students’ overall learning.

”

7.3.2 Complementarity and Agreement between Formal and Informal Sources

In spite of their dependence on class tests (at least, as far as formal assessment is concerned), the participants demonstrated awareness that assessment through class tests can be problematic. In fact, many of them warned against viewing test marks in isolation, and argued instead that a credible interpretation of test results requires a good knowledge of the underlying context (see Smith, 1991; Torrance, 1995). Table 7.4 summarises their reasons for regarding test marks with utmost caution.

Table 7.4: Teachers' reasons as to why test marks may be misleading

- **Student factors:** Students who fear tests are not likely to show what they know under examination conditions. Moreover, should they either panic during a test, or be passing through a stressful period in their lives, or suffer from poor memory, or not prepare adequately for it, underachievement is also likely to follow. On the other hand, students may do better than they deserve should they either chance on questions for which they are particularly prepared, or copy during the test.
- **Teacher factors:** The difficulty level of a test varies from test to test, and from teacher to teacher. In a relatively easy test, the student marks are likely to be generally higher than had the test been more demanding. Again, the amount of pre-test hints and 'guidance' offered by the teacher is also likely to influence the final outcome. The result is also affected by the marking criteria adopted by the teacher.

Table 7.4 reveals that most of their misgivings are not directed at the tests themselves, but rather at the possibility of uninformed judgements being drawn from their results. At the same time, each teacher presented him or herself as the most competent person to interpret his or her own test results by virtue of being singularly knowledgeable – particularly due to the ongoing assessment information to which he or she is daily exposed – about the embedding teaching-learning situation (see Calfee & Masuda, 1997). The participants thus view the teacher as that person who can best position the test results within the much wider classroom assessment context from which the marks draw meaning and to which they add meaning. This suggests that whilst tests provide teachers with a time efficient tool to gather information about students (see Denvir, 1989), teachers still feel the need to invest in time and reflection when they come to interpret what lies behind the test results. Andrew crystallised this argument:

When I see the test mark, I look beyond the mark ... I visualise the student concerned and think about what he has been doing since he came to my class ... I also keep in mind what kind of test I gave, what I was expecting from him, whether I corrected in a lenient or strict manner ... Only because of this background can I really talk about the mark ... There are so many things the mark doesn't show about the student, but which I know ... But then the mark helps me also to understand whether my impressions about the student are correct ... I see what I learn from tests and the classroom as two sides of the same coin.

This 'reflective exercise' mirrors in turn the complementarity noted by the teachers between the classroom assessment results obtained from tests (or formal assessment) and other sources (or informal assessment). Teachers claimed in fact that their understanding of the teaching-learning situation deepens by combining insights from both formal and informal sources (see Airasian, 2000), which they largely view as 'pulling in the same direction'. So much so, that teachers repeatedly stressed that, except for the 'student factors' listed in Table 7.4 (which they see as affecting only a small minority of students, and not all the time), their assessment results from class tests (read 'formal') and other sources (read 'informal') practically reflect each other.

But the teachers usually qualified that this 'agreement' between the results of formal and informal assessment practices needs time to materialise. They explained that they start to learn about their teaching-learning situation in class – which in Airasian's (2000) terminology is 'sizing up assessment' – through everyday informal assessment sources, as they only resort to formal assessment practices after a number of weeks (if not months) have already passed. The teachers' initial complete dependence on informal assessment reportedly lasts until the onset of the first class test. In fact, they explained that, from then onwards, a sort of 'knowledge interplay' sets in between what they learn from their formal and informal practices. They added that this process continues until the end of the year with both sets of information sustaining each other by reciprocally adding definition and meaning on an ongoing basis. This is how Nicholas described these progressive and mutually redefining exchanges between his formal (i.e., class tests) and informal assessment practices:

In the first few weeks, I start building ideas about individual students from what I see and hear in class ... at first, these are still vague impressions! Then the first test gives me the chance to reconsider these impressions – I may either confirm them, or change them slightly, or even abandon them completely! ... Following the first test, the whole process starts all over again ... once again I continue shaping my ideas according to what goes on in class, and this until the next test either

confirms or challenges them! ... But as time passes, my assessment ideas become more stable and ever less likely to change with each new assessment result.

Nicholas' reference to the stabilisation of his assessment 'ideas' over time – a process through which all the other participants also reportedly pass, and which somewhat contrasts with Airasian's (2000) assertion that 'sizing up' impressions tend to become permanent – signals the strong belief amongst them that the gradual accumulation of information reaching them from a variety of sources eventually provides them with a sound knowledge of students as far as mathematics is concerned (see section 7.6.1).

“

Michael's Story

Like most colleagues, tests are my only direct source of student marks. Believing that test scores need to be interpreted from a knowledgeable contextual position, I share my colleagues' misgivings about the multiple dangers of drawing uninformed conclusions from test results. I also concur with their view that the class teacher, being so informed, is the most suited person (I would however add 'together with his or her students') to interpret test results. Information from my formal and informal sources, especially over a time period, often points in the same direction. But rather than being appeased with this general 'agreement', I have learned that there is much more to gain by unearthing and investigating 'disagreements'. This exercise provides me with a richer assessment picture that makes it more possible for me to promote learning with my students.

”

7.3.3 The Use of Formal and Informal Sources for Official Reporting

The teachers report officially on their students at two different levels: (i) when an official request (say, by the school administration or a parent), albeit very rarely, is made; and (ii) when they give the statutory Assessment Marks. Just as they link formal sources with 'official' assessment and informal sources with 'unofficial' assessment (see section 7.3.1), so they associate the results of formal practices (particularly from class tests) with an official way of communicating assessment results and the results of informal practices with an unofficial way of communicating assessment results. Nicholas highlighted this distinction when he said:

If a parent would want to know about his son or daughter, I'll just talk about the marks ... I'll give him something concrete! Unless they ask me, I won't go into

what goes on in class ... the marks are official and I'm bound by them! But if you were to ask me, say, about your cousin, I wouldn't mind telling you whether or not he participates in class, if he pays attention, if he does his homework, if he makes an effort or couldn't care less ... I would give you this other picture. But at the official level, I'll prefer to stick to the tests!

Indeed, although the teachers are open to receiving feedback from a variety of sources (see section 7.2.2), I have noted a certain restraint on their part when it comes to communicating feedback at an official level that is not strictly based on marks. The strong link between official reporting and the numerical records of formal assessment (see section 7.5.1) is mostly evident in their grading practices. The teachers reported in fact that what they learn from informal assessment is either completely ignored (two of them mentioned this), or else is only used to round marginally up or down the Assessment Marks (which, in any case, have to be presented to the school administration in multiples of five) as they see fit. This grading 'bias' against knowledge gained through informal assessment (see Denvir, 1989) shows that at the most 'official' level of communicating the results of classroom assessment, it is formal assessment that actually matters (see Airasian, 2000). Matthew provided a typical account of how the Assessment Marks are computed:

I normally give them two tests each term ... I start by calculating these tests' average for each student ... Let's say a student gets an average of 60 ... At this point I ask myself, "How is he doing in class? Does he ask questions during tutorials? What do I know about him?" ... Things like that! If I find that my general idea about him matches his average of 60, which is generally the case, he'll get an Assessment Mark of 60. However, if I think he deserves either more or less than 60, I slightly increase or decrease the mark of 60 accordingly ... but these changes are only minor ones!

Ray is the only teacher to depart significantly from this procedure. He reportedly allots 20% of his Assessment Marks to attendance, as he thinks it is very important that 'students are exposed to mathematics and be credited for this'. But then, like all the others, he only uses the results of informal assessment for rounding off purposes.

The teachers offered two explanations for this 'rounding off' that, as the evidence reveals, alters the formal assessment results (and 'attendance' in Ray's case) between plus or minus 10% (even if teachers normally round up the marks, not down). First, a few teachers claimed that they occasionally utilise rounding as a psychological tool in order either to reward and encourage the hard working but mathematically weak

students (this involves rounding up – see Airasian, 2000), or else to keep the very high performing students from becoming overconfident and consequently slacken their efforts (this involves rounding down). Second, most teachers claimed to use rounding in order to partially bridge any arising differences between the results of formal and informal assessments. They argued that such ‘corrective interventions’, if and when they become necessary, guarantee a greater degree of fairness in the computation of the Assessment Marks – a way to ensure that, in most cases, the ensuing grades are a concise representation of what the students know in a manner that everyone can easily understand. But even though the teachers feel the need to use their ‘professional judgement’ on grades (see Loyd & Loyd, 1997), this form of personal discretion is only applied very sparingly. Rita explained why:

What would people say if my Assessment Marks are not close to my test and homework results? I can explain little discrepancies by saying, for instance, that the student has been trying hard or is unmotivated ... But as our system is based on marks, everyone expects me to base Assessment Marks on marked work!

Rita’s concern to base her official results on student marks so that they may be easily defended in public (see Airasian, 2000; Peterson & Stack, 1998) emerged as a top priority amongst all the participants (see also section 7.5.3). In fact, all the teachers claimed that whenever the differences between their personal evaluations and the written assessment records cannot be easily reconciled through normal ‘leeway measures’ (even if this, according to them, rarely happens – see section 7.3.2), they always stick much closer to the results of formal assessment. And as their formal assessments are overwhelmingly class tests, the Assessment Marks are consequently by and large a reflection of students’ performance in these tests – a sort of a class tests’ average mark within a particular term. This ‘link’ is again evident from the low weighting given to non-test marks (about which teachers have serious authenticity doubts – see section 7.2.2.2) in comparison to test marks (even if teachers are aware that test results may be misleading – see Table 7.4) by two of the three teachers who have student marks on work other than tests (see section 7.2.2.2). Given that all the participants practically base their Assessment Marks on ‘checking up’ rather than ‘summing up’ assessments (see Harlen et al., 1992), there appears to be – contrary to what they think and fear (see section 7.1.1) – some sort of undisclosed, baseline criteria governing the role of classroom assessment in the school’s promotion system.

“

Michael's Story

I compute the Assessment Marks in largely the same manner described by my colleagues. I start by averaging the test marks obtained during a particular term (i.e., if more than one test was set), and then proceed to adjust these averages (by up to plus or minus 15%) according to my knowledge of students from informal assessment. Then, as I round off the resulting marks to the nearest 5% (as stipulated by the school regulations), I tend to make slight allowances in students' favour for regular attendance, participation in class, evidence of interest and willingness to learn, and judicious use of personal contact hours. Notwithstanding the considerable weighting I give to test results in my Assessment Marks, I am under no illusion that these marks provide a good indication of what my students actually know and can do. I see them instead for what they are – only rough estimates of what I perceived that my students were willing to do and reveal during a certain period of time, which no 'pseudoprecision' exercise on my part (see Anderson, 2003) can ever make more exact. This explains why I prefer to add what 'I know' – even if there is the risk that this may be perceived as 'mere impressions' – on the very rare occasions when I receive an official request for information about a student. I find that although people, especially parents, are greatly interested in 'performance', once I gain their trust and attention, they become very eager to hear about the fuller picture that I can provide.

”

7.4 The Classroom Assessment Tasks

7.4.1 The Centrality of Tasks in Classroom Assessment

Mathematical questions (or tasks) are an omnipresent feature in PMI lessons. The participants, in fact, normally dedicate a substantial part of the exposition phase of the lesson to working examples on the whiteboard, the students are expected to work questions during the practice phase, and dealing with students' difficulties arising from working questions practically accounts for the consolidation phase (see section 6.5.3). This is how Mario referred to this reality:

Doing mathematics is about working questions ... all my teaching efforts are meant to help students do just that! ... Questions are also part and parcel of classroom assessment ... Aren't class tests about working questions? Don't most of student difficulties concern questions? And what do I see when I check students' work? Questions! And what about what I ask them in class? ... Wherever you look, it's all about questions ... directly or not!

Tasks – around which TCAP gravitate – thus constitute a vital means for teachers to learn about their teaching-learning situation. The importance of tasks is reflected in

turn in their general belief that all is well if a student can do a task, but that he or she (i.e., should the student be one in whom it is considered worth ‘investing’ – see section 6.4) has to be re-taught if this is not the case.

“

Michael’s Story

Tasks are just as central to my assessment practices as they are to my colleagues, and in pretty much the same manner. But then, I do not share the other participants’ belief that ‘everything is fine once the task has been worked correctly’, as I know that students can perform well on tasks that they do not even understand (see Desforges & Cockburn, 1987).

”

7.4.2 The Criteria for Selecting Classroom Tasks

To explore the participants’ criteria for selecting mathematics questions, I presented them with a set of four tasks that represent varying degrees of ‘routineness’ and demand varying levels of thinking from students (see section 5.3.2.2 and Appendix II). The teachers’ comments on these four tasks (see Table 7.5) left no doubt that they overwhelmingly favour the use of traditional, routine tasks (read ‘closed tasks’ – see Boaler, 1998), and that the further a task moves away from this ‘low thinking level modality’, the lesser is the likelihood that they would want to use it (see Carter & Doyle, 1987). So much so that the teachers, contrary to their universal acceptance of the very traditional Task 1, largely dismissed the use of the non-traditional Tasks 3 and 4, albeit being generally very appreciative of their educational value, and spoke of the need to ‘redefine or simplify’ (see Doyle, 1988) Task 2 by stripping it of certain ‘parts’ to render it less nonroutine, and thus more acceptable.

When most of the teachers grouped the four given tasks into those that are similar to the ones they use (i.e., Tasks 1 and 2) and those that are not (i.e., Tasks 3 and 4), they were not making a statement about what they see as ‘ideal’, but were distinguishing instead between tasks that ‘fit the system’ (or ‘could be made to fit the system’) and others that do not (see Buhagiar, 2004). This reveals the extent to which their task choice is context-bound (see Broadfoot, 1996). In particular, three levels of context –

Table 7.5: Teachers' comments on the four tasks

- **Task 1:** In comparison to the others, the teachers spoke very briefly about this task – showing more ‘acceptance’ than ‘enthusiasm’ in the process. Describing it as ‘typical’, ‘normal’, ‘standard’, ‘traditional’, and even ‘classic’, they concurred that this task has wide classroom application, such as classwork, homework and class tests. The teachers also agreed that Task 1 is a worthwhile activity as it: (i) is the method type (i.e., once the students identify the correct method, the solution follows without difficulty); (ii) is well structured (i.e., the task structure itself guides the students towards the solution); and (iii) serves as good practice (being so typical of both textbook and examination questions) to prepare for the examination.
- **Task 2:** The teachers generally agreed that this task could also be used, especially if some of its parts are either adapted or eliminated. In particular, they suggested that the task would become more suitable for classroom use should it ask students directly to minimise the surface area, and should it not require students to state their ‘assumptions’ and to discuss the ‘shape of the can’ (things that, in their view, could still be discussed informally in class). The teachers pointed out that, unlike Task 2, examination questions are usually very direct (thus avoiding long descriptions in English that most Maltese students would have difficulty understanding) and do not require students either to state assumptions or to analyse shapes. A few teachers commented further that albeit Task 2 appreciably encourages students to think instead to merely recall and use ‘past’ information, they would still not use it without modifications, as it would be ‘too much’ for their students.
- **Task 3:** Most teachers acclaimed Task 3 as an activity that helps students to develop their mathematical thinking and teachers to gain insights into how students think. Some of them, however, expressed doubts about students’ ability to tackle such descriptive tasks in English. As they read through this task and started considering the various ways in which it can be worked out, some teachers commented that this was a ‘real’ challenge – not only for their students, but also for themselves. They claimed to lack the class time needed to tackle such challenging tasks. Thus, in spite of its many declared merits, the teachers seemed unwilling to use Task 3 widely (if at all) in their classrooms. They said that, at best, they would consider using similar tasks very occasionally, and only as basis for teacher-led classroom discussions.
- **Task 4:** Most teachers described Task 4 as an interesting and challenging activity (even if not as much as Task 3) that stimulates students to think mathematically. Notwithstanding this, only a few of them said that they would consider using similar tasks in class, and then only very occasionally. These few agreed further that should Task 4 be used in class, it would have to be during a teacher-led classroom discussion. Two teachers argued that the ‘missing measurements’ should be added to the task, as students would otherwise have difficulty dealing with it. One teacher refused to accept this task as a proper mathematical task – he argued that it is more like a ‘game’ than a problem.

namely, the national, school and personal contexts – were identified that influence, both on their own and interactively, the way in which the teachers select tasks. Table 7.6 details the teachers’ task selection criteria by these three contexts. Of all the criteria listed in Table 7.6, the one that appears to carry the major weight in teachers’ decision is that the task should mirror MATSEC examination questions (the first year teachers reported that although their students sit for the school’s end-of-first-year examination, they still follow MATSEC questions as the school examination ‘is

known' to be based on them). This reality of given attention in class to what gets examined (see Gipps & Stobart, 1993) was succinctly exposed by Stephen:

One has to be practical! ... As long as MATSEC keeps dishing out examination questions like Tasks 1 and 2, that's what I'll be doing in class ... what I think about them, doesn't really have to matter!

Table 7.6: The task selection criteria by the three determining contexts

- **The national context:** All the teachers insisted that the tasks must follow MATSEC examination questions (which, incidentally, have been judged by most teachers as increasingly pitching at the right level). It is also crucial for most teachers that the selected tasks not only do not lead to syllabus coverage difficulties, but that they also reflect the current secondary level mathematics programme (even if this does not, in their view, provide an adequate preparation for post-secondary level mathematics – see section 6.3.5).
- **The school context:** Many teachers claimed that when choosing tasks they also keep in mind that the school does not operate a selective entry policy (see section 2.4.2). These teachers hold that, as a result, many students still have gross difficulties with basic mathematics and English (which is the tasks' written medium). They thus try to avoid 'too demanding' and 'too verbose' tasks. Many teachers also claimed to avoid tasks that 'require too much time', as it would not be otherwise feasible to cover the 'vast' syllabus content with just three lessons per week.
- **The personal context:** Judging by what many teachers said, tasks are also chosen because they mirror their views on the nature of mathematics (read 'as a fixed, static body of knowledge' – see Romberg & Kaput, 1999), as well as appertain to the type of mathematical problems they have been working with since their student days. A couple of teachers added that as the PMI option is 'pure' (as different from 'applied'), the tasks should not present the application of theory in practical situations.

Given that the participants' test questions follow closely those used during lessons (see section 7.2.2.5), it is not hard to understand that their tests mimic the MATSEC external ones (see Bright & Joyner, 1998; Mavrommatis, 1997; McCallum et al., 1993; Torrance & Pryor, 1998). The 'link' between class and examination tasks explains the teachers' oft-repeated claim that they regularly examine MATSEC examination papers in detail (see section 6.2) – a scrutiny exercise that reportedly helps them to reconfirm (or refute) the existing task patterns and to establish new ones. This analysis helps teachers to decide which questions to use in class, as they are most unwilling to devote class time to questions that they deem irrelevant to the examination (see Goldin, 1992). Many teachers argued moreover that it is particularly worth analysing the MATSEC question patterns in view of the continually falling student standards (see

section 6.3.5). Angelo highlighted their reasoning when he said, “*What’s the point of trying to extend our students beyond the bare minimum, when so many of them have to struggle even with that?*” These teachers generally explained their ‘minimal’ position – which somewhat echoes LaCelle-Peterson’s (2000) argument that assessment is used to legitimate minimal services to students – by saying that it denotes a ‘sign of respect’ towards students, as it increases their likelihood of eventually passing the examination (which is, after all, what the teachers reportedly desire most for their students and what they think the students expect from them – see section 6.4).

“

Michael’s Story

Initially, I stuck primarily to the tasks of the official textbook, particularly those recommended by the more experienced colleagues (Nicholas was particularly helpful in this respect during the first year). But as my understanding of what I-Level pure mathematics involves grew and I gained in self-confidence, I started to look beyond the official textbook. I was at that point picking tasks (some of which I still use) from various textbooks – my main selection criteria being that the tasks fitted the PMI syllabus, had the potential to ‘develop learning’ rather than just ‘produce work’ (see Desforges & Cockburn, 1987), and were varied and time efficient (i.e., except for the few more open tasks). I paid little attention to the MATSEC (and the school) examination questions until then – and this according to my belief that, more than preparing students for a specific examination (albeit I still aim for syllabus coverage), teaching is about helping students to learn mathematics. But then, I began to realise that in spite of my emphasis on ‘learning’, my students were not achieving as good MATSEC examination grades as was to be expected. In the knowledge that students need a pass grade or more to join certain university courses (see section 2.3.2), I decided to adopt a more pragmatic approach than my beliefs would normally allow me – albeit knowing that this would benefit students only in the short term (see Torrance 1995; also Shepard, 2001 [cited in Schoenfeld, 2002]). This is how I began to emphasise in class, particularly towards the end of the year, the MATSEC examination type of questions. And although I still have misgivings about doing this, I think that my students’ chances of making it to university have certainly improved as a consequence. This shows that the criteria I use to select classroom tasks mirror the same contexts that I identified for the participants – namely, the national (e.g., paying special attention, as of late, to MATSEC type questions), school (e.g., choosing time efficient tasks), and personal contexts (e.g., still using non-MATSEC type questions).

”

7.4.3 The Nature of Classroom Tasks

To understand better the nature of the MATSEC examination inspired paper-and-pencil tasks used in class, apart from the interview data, I also reviewed the teachers’

class tests and worksheets used during Year 2 of the study, the textbook exercises that they reportedly used in class, and the PMI MATSEC examination questions. The characteristics elaborated in Table 7.7 – which gives the salient points of this task analysis – indicate the pervasive use within PMI classrooms of lower order tasks that do not normally require students to go beyond mere factual recall of procedures learned previously in class (see Black & Wiliam, 1998). In what appears to be a long

Table 7.7: The nature of classroom tasks

- **Close-ended:** The task's solution is such that only one specific answer is correct. In fact, the teachers who suggested that 'measurements' should be added to Task 4 were basically trying to redefine this task from an open-ended activity (i.e., all reasonable answers are judged to be correct) to a close-ended one.
- **Highly scaffolded:** The task is closed-middle, meaning that its very structure leads the student towards the solution through known algorithms and procedures. Basically, all the student has to do is to establish some well-known solution method (from a range of class-taught ones) and follow it.
- **Lower order:** The closed nature of the task (being both close-ended and highly scaffolded) renders it a lower level thinking activity. Contrary to the challenging intellectual demands of the more open mathematical tasks (such as practical and investigative work), real thinking is here reduced to a minimum – with solutions relying primarily on the application of recall procedures (e.g., substituting values in a formula).
- **Topic specific:** The task can be readily identified within a specific syllabus topic (e.g., 'solving trigonometric equations' as part of trigonometry). And if more than one topic is involved, the link between these topics would have already been discussed and firmly established during the lessons (e.g., the link between partial fractions and integration in Task 1).
- **Direct and minimal use of language:** The wording of the task, as evidenced by the teachers' insistence that students should have been asked specifically to minimise the surface area in Task 2, is generally very straightforward. Moreover, long descriptive language is avoided in tasks – suffice it to point out that this was one of the reasons as to why the teachers did not consider Task 3 to be appropriate for wide classroom use.
- **Time efficient:** A task must not be too time consuming. As it is rather 'easy' to understand, it does not require too much time to determine an appropriate solution method and to execute it. The lack of conformity of Task 3 with this criterion was another reason behind its failure to be looked favourably upon by the teachers for wide classroom use.
- **Strong theoretical bias:** The presentation of the task is orthodox and very mathematical (in fact, a few teachers partially explained their reluctance to use Task 4 as it fails the traditional canons of mathematical problems by not specifying 'measurements'). Again, a task is primarily intended to examine whether or not students have competence at theoretical level. In fact, the task hardly ever integrates theory with its application to real life situations.

established practice, the teachers use very traditional activities (read ‘closed, routine, examination type tasks’) that, by their own admission, some of them do not even think too highly of. These tasks: (i) do not take too much of teachers’ time; (ii) hardly cause disruption in class as students can work on their own without too much trouble; and (iii) serve as good practice for the school and MATSEC examinations that await students at the end of the first year and the second year of the PMI course respectively. But these tasks – which have little chance to develop mathematical understanding and to reveal students’ thinking – almost certainly fail what Boaler (1997) identifies as the most important aim for the teacher, that is, to engage students and to provide worthwhile activities that they find stimulating (see also Desforges & Cockburn, 1987). For this is highly unlikely to materialise with tasks, such as the ones used by the participants, that do not require students to formulate problems and develop and apply strategies to find solutions in a range of contexts, and then to verify and interpret results and generalise solutions – an exercise that would undoubtedly serve students’ needs to make sense of experience arising outside of mathematics instruction and mathematics itself (see Romberg & Kaput, 1999).

“

Michael’s Story

The characteristics of my classroom tasks also largely follow those detailed in Table 7.7. The only notable exceptions being when I very occasionally present my students with a relatively open task in order to get them involved in a ‘real’ classroom discussion (see section 7.2.2.4). My quite recent ‘rapprochement’ with MATSEC type questions has thus not put an end to my by now established, albeit very occasional, forays with relatively open tasks. Indeed, the only real consequence of my greater attention to MATSEC questions is that when deciding between equally closed and traditional questions, I now stick closer to what appears to be ‘more in fashion’ with the MATSEC examiners.

”

7.5 The Results of Classroom Assessment

7.5.1 Recording the Results

Similar to the ‘intuitives’ in Gipps et al.’s (1995) model, the participants have very little written assessment records. Indeed, other than Kathleen, they only have numeric

records (or ‘marks’ that resemble the summative records mentioned by Murphy and Torrance [1988]) from class tests (all the teachers have them) and homework (just three teachers have them). And from what Kathleen said, her non-numeric records appear to be rather concise and sporadic:

Other than marks, I sometimes write short notes like ‘this student has difficulty with logarithms’ or ‘I should avoid using this example in the future, as it’s too difficult for them’ ... just things like this that I keep with my records. I don’t write these notes regularly ... only when I happen to remember and find some time.

Apart from these ‘permanent’ assessment records, some teachers reportedly also keep temporary ones. For instance, Mario maintained that after each class test he makes a detailed list of student mistakes so that he could draw his students’ attention to them later on in class. These records are only kept for a very short time – in Mario’s case, until the test is discussed in class. Notwithstanding the existence of these temporary records, the bulk of what the teachers learn from their assessment practices is still only being recorded mentally, not in writing. This is even more so in the case of assessment results originating from informal practices, which are by far more likely to produce evidence in ephemeral rather than permanent form (see William & Black, 1996). Angelo referred to this reality when, with reference to his informal practices, he said, “*I just store in mind what I learn ... I certainly don’t keep written dossiers on students*”. All the teachers however reassured me that their myriad of mental records, together with the few written ones (which they use primarily to compute the Assessment Marks – see section 7.3.3), more than serve their classroom requirements.

“

Michael’s Story

Next to the names of students who come to visit me for personal contact hours, I write short notes (comparable to Kathleen’s non-numeric records) in which I describe briefly the nature of the difficulties discussed (e.g., The student had difficulty with trigonometric equations, particularly those requiring the use of the Pythagorean identities) and the outcome of the encounter (e.g., Advised student to work exercise so and so, and to get back to me next week). But other than this, my assessment records are as ‘numeric’ and ‘mental’ as those of my colleagues.

”

7.5.2 Responding to the Results

Whenever the teachers referred to ‘classroom assessment’, the underlying idea was consistently that of obtaining information on their teaching-learning situation as perceived within the prevailing ‘teaching as transmission and learning as practice’ environment (see section 6.5.3). Within this scenario, classroom assessment emerges as the means for the participants to determine the effectiveness with which the knowledge they are trying to teach has been communicated to their students (see Brown, 1989). Andrew put it like this:

Classroom assessment is like a thermometer! I use it to check my ‘patients’ and myself. It informs me if my ‘message’ is reaching the students or not, whether they are following me or not ... well, at least, most of them ... As you well know, it is often the case that the teacher does not even recognise his own ‘message’ once it gets back to him through some feedback or other!

Revealing an understanding that action should follow the outcomes of assessment (see Wiliam & Black, 1996), all the participants concurred that it is the teacher’s duty to respond in some form or other to what the results of this ‘effectiveness check’ would be indicating. The manner in which they reportedly respond may be classified under these four categories: (i) work of consolidation; (ii) advice to students; (iii) changes in the teaching approach; and (iv) changes in classroom assessment practices.

The **consolidation work** – which is one of the three pillars of the typical PMI lesson (see section 6.5.3) – refers to those episodes (which were mentioned, in some form or another, by all the participants) during which the teacher sets out to clarify students’ difficulties that would have been identified either by the student or students concerned, or by the teacher him or herself. This consolidation occurs during lessons (which is often the case), tutorials (should this apply), and personal contact hours. Table 7.8 identifies the three sources of students’ difficulties and details how the teachers go about consolidating student’s learning in each particular case. Their consolidation work, which is inspired by the traditional transmission model in which students are passive recipients of knowledge, appears to be largely geared at ‘unblocking’ whatever might be keeping students from receiving knowledge (see Denvir, 1989).

Table 7.8: Teachers' work of consolidation

- **Exposition:** Whilst the teachers are presenting new theory and solution methods, the students may encounter difficulties and make these known. The teachers reported that, faced with such difficulties, they usually end up 'repeating' what they would have just said or done in an effort, as Andrew put it, "... to get the message through the second or third time round". However, the teachers also remarked that at times it becomes necessary to go beyond mere repetition. This involves rephrasing things, adding or reducing details, and making use of other examples.
- **Practice work:** The students practise solution methods by working exercises given by the teacher for classwork and/or homework (and at a later stage by working examination past papers). The teachers reported that when students face them with practice work difficulties (which usually occurs at the start of the lesson or during tutorials), they normally end up working out the indicated question or questions for the whole class. The teachers pointed out that during the lesson they lack the necessary time to go individually over students' practice work difficulties. However, they all agreed to have more time for this during tutorials (where this applies) and/or personal contact hours.
- **Class tests:** Correcting class tests reportedly offers teachers the opportunity to note (and in some cases also to take note of) students' mistakes, particularly the more common ones. Most teachers claimed that after giving the corrected scripts and marks to students, they normally work out the test paper in class. This gives them the chance to 'point out' and 'tackle' the mistakes encountered. The few teachers who do not do this class correction explained that as the students are primarily interested in marks, they see no point in 'wasting' a whole lesson on it. These teachers prefer instead to pass some very generic comments to the whole class concerning their performance in the test that normally only take a couple of minutes.

Another much diffused teacher response to the results of classroom assessment is to offer **advice to students**. The teachers explained that this advice (usually in the form of a phrase uttered publicly inside the classroom) serves two main purposes: (i) to encourage those students who are working hard and making good progress to keep up their efforts; and (ii) to warn those who are doing neither of these to change their ways (the more frequent recommendations being 'to pay attention in class', 'to do the set work', and 'to study'). Either way, their advice appears to lack the 'gap-closing' information that helps students to improve (see Ramaprasad, 1983; Sadler, 1989; Wiliam & Black, 1996; also Bright & Joyner, 1998; Brookhart, 1999). The teachers moreover admitted that they find it hard to give advice to students who show no significant progress (and in some cases, even perform very poorly) in spite of making a good effort. As Rita explained, the inability to pinpoint the blame on something 'safe' (e.g., 'not enough practice work') puts the teacher in a rather awkward position:

I just don't know what to say to students who keep getting poor results in spite of working hard ... if I couldn't blame their efforts, what could I tell them? I won't ever tell them that they're not meant for mathematics, even though it may be so!

The teachers remarked also that when their advice is based on the results of formal assessment (especially, tests marks), the students are more likely to take it seriously and act accordingly (see Denvir, 1989). On the other hand, they largely agree that with students who ‘refuse to learn’ and consequently ‘cannot be bothered about anything’, no amount or type of advice they could ever give them would make any difference. Teachers consequently prefer to ignore such students as long as their classroom behaviour does not impact negatively on the flow of the lessons (see section 7.2.2.1).

Making **changes in one’s teaching approach** is another form of teacher response to the results of classroom assessment. Reportedly, these changes have invariably to do with the teacher trying to match his or her teaching ‘level’ according to the perceived quality of the students. In general, it is more or less a question of finding the ‘right’ teaching level to a particular class or student (even if only a couple of participants mentioned this individualised form of response), and subsequently acting upon it either within the same lesson or in the following lessons. Andrew offered a typical account of this widespread phenomenon:

It is through assessment that I understand what level of students I have and how it is therefore best to teach them ... Assessment feedback informs me if my message is reaching the students or not, whether they are following me or not ... Then, according to what I learn from it, I either say, “The lesson was OK ... the message got through” or “It flopped ... the message didn’t get through”. When the feedback is positive I may, for example, give them more challenging examples next time round or start skipping some easy steps in my working ... But when the feedback is not positive, I would start thinking whether I should extend the introduction, or move at a slower pace, or include more worked examples ... things like that! ... Sometimes it’s just not possible to make any changes during the lesson itself ... these would then have to wait for the next time I meet them!

In line with Wood’s (1990) observation that teachers see test results in terms of students rather than themselves, the participants primarily view the results of assessment as a reflection on the quality of their students, certainly not on their teaching method. In fact, when they referred to how they ‘change their teaching approach’ in accordance to the assessment feedback received, it was always a question of ‘level’ (e.g., giving more practice questions when the practice work results are still considered to be unsatisfactory), never of ‘method’. Indeed, their teaching method reportedly remains unchanged (practically entrenched, once established – see section 6.3.2) irrespective of how negative the feedback received may be.

A number of teachers also mentioned affecting **changes to their classroom assessment practices** in response to assessment results. Whilst some of these changes remain localised to a particular classroom or student situation (e.g., Renzo said that he does not question directly the very weak students, as this would halt the smooth flow of the lesson – see section 7.2.2.4), others changes go beyond this and carry longer term consequences (e.g., Andrew reported that his class tests are becoming increasingly less demanding in view of the continually ‘deteriorating quality’ of the students – see section 6.3.5). All the reported changes in teachers’ assessment practices moreover concern matters that are essentially related either to ‘quantity’ (e.g., less and less teachers are collecting, correcting and marking students’ – see section 7.2.2.2) or ‘level’ (e.g., teachers tend to discard highly descriptive assessment tasks in view of the students’ perceived difficulties with the English language – see Table 7.5). In fact, except for Ray’s ‘experimentation’ with class tests (see section 7.2.2.5), there is no evidence to suggest that such changes alter the underlying nature of teachers’ assessment practices in any significant way, not even remotely.

“

Michael’s Story

*Whatever I might learn about my teaching-learning situation would be of little use unless I am then able to respond (and to help others respond) in a manner that brings about improvement. This is best achieved when teaching, learning and assessment are well integrated so that the classroom environment supports rather than inhibits learning (see Black, 1998; NCTM, 1995; Torrance, 1995). I find however that my responsive action (which can also be classified under the participants’ four categories) is greatly curtailed both by the nature of what I come to know and by the operating context. For instance, although I give a lot of importance to **consolidation work** both inside and outside the classroom, I am ‘painfully’ aware that my efforts are exceedingly of the ‘directly solving students’ difficulties’ type as opposed to the more learning-friendly ‘helping students to solve their difficulties’ type for reasons that primarily have to do with time constraints (particularly inside the classroom).*

*In my quest to offer **advice to students**, I find it rather helpful to know what they are hoping to achieve by studying the PMI option. This provides me with a sense of focus as I encourage those who are doing well to keep up their good work and give the ‘wake up call’ to those who are starting to fall behind. I do admit however that with some students (especially those who were pushed into the option by the matriculation regulations [see section 2.2.1.2] and those who would have lost all interest in it in the meantime) no amount of advice ever seems to make any difference. For although they may listen to me politely, and possibly admit to their faults and make well-meaning promises, they are usually back to their ‘old selves’ before long – which explains in turn why I find it ever so hard not to give up on and consequently ignore such students.*

All received classroom assessment information has contributed to ‘changes’ in both my teaching approach and classroom assessment practices. But, in either case, irrespective of whether these changes developed over the years or are in direct response to a particular classroom situation, they have never as yet altered or seriously challenged the very nature of my teaching or assessment practices. Indeed, although I continually affect modifications to my teaching approach in the light of the incoming assessment information, my teaching remains essentially transmission in nature (see section 6.5.3). The same applies to my assessment practices. For, like my colleagues, it has been mainly a question of gauging the ‘right quantity or level’ according to the arising situations inside the classroom. So far, in reality, all that I have to show for ‘visible assessment change’ are my recent little ‘excursions’ away from the conventional testing formula (see section 7.2.2.5).

”

7.5.3 Communicating the Results

The teachers reportedly offer ongoing ‘unofficial’ feedback to students based on their informal assessment practices. Such feedback, direct or indirect as it may be, can take the form of a simple nod of the head, an encouraging remark, a rebuke, a clarification, and so on – all things that I actually observed inside their classrooms. At the same time, their holding back from officially divulging feedback that is based on informal practices (see section 7.3.3) appears to result from a ‘nagging feeling’ on their part that persons on the receiving end do not take such feedback as seriously as that that originates from formal practices. The teachers apparently fear that feedback based on informal practices would not be given much credence (even by colleagues), as it is not supported by ‘marks’ that are largely believed to measure what students learn (see Gipps, 1994). The following passage from the fieldwork journal illustrates this reality:

Nicholas has been complaining for some time about his PMI class. Apparently, it is one of the worst classes he has ever taught. ... This morning he entered the office with a strange grin on his face, and after he almost ‘hurled’ at our faces the results of the test he gave the other day, he exclaimed with satisfaction, “Now you can see how right I am ... only five passed! ... Look at the results! ... Eight of them did not even get more than 10% ... this proves that I have been correct all along!” (November, Year 2)

It may thus be that, albeit teachers’ formal practices undoubtedly produce what Broadfoot (1996) calls ‘dead data for reporting purposes’ (see section 7.5.1), they also serve to lend ‘credibility’ to teachers. The above anecdote highlights the participants’ manifest uneasiness about not being able to externalise convincingly what they learn from informal practices. Given that this type of information reportedly constitutes a

very significant portion of their assessment knowledge (see section 7.3.1), the teachers may well be keeping to themselves, particularly at the official level, a substantial part of what they actually learn about their teaching-learning situation.

Very often, the only ‘official’ feedback that the teachers pass on to their students is the marks. Many teachers referred to these occasions, particularly when test marks are involved, as ‘moments of truth’ for students – times when students are made unequivocally aware of where they stand. A few teachers – in line with the participants’ norm-referenced understanding of assessment (see Black, 1998) – reported to supplement the individual test marks with some additional test statistics, such as the mean and the range of marks, so that the students would be in a better position to evaluate their performance in relation to others. Sometimes, both test and homework marks are accompanied by brief comments. For instance, Matthew said that above and below a certain test mark, which may vary from test to test according to the level of difficulty involved (see Table 7.4), he normally writes on the scripts comments like ‘keep up the good effort’ and ‘you haven’t really studied’ respectively.

But such comments (which are often very generic) hardly ever make it to outside the classroom confines. The only tangible evidence of classroom assessment results to reach here is the school’s assessment report sheet that teachers have to fill in whenever they grade students. For every student, this report only includes the Assessment Mark and any comment that the teacher may wish to add inside a rectangular space with dimensions circa 1 cm by 4 cm. According to the subject coordinator, who collects these report sheets before passing them on to the school administration, most teachers leave this space blank – an observation that was also corroborated by the teachers. Rita succinctly highlighted the participants’ reasons for this:

Do you think the administrators are actually interested in reading our remarks? All they’re interested in is the mark ... that helps them to determine who’ll be promoted or not! ... And anyway, there isn’t much space in which to write either!

The teachers appear reluctant to communicate assessment information that seems to serve no real purpose. A person from the school administration office informed me that these teacher comments, apart from being largely ignored by the administration, are not included in the assessment reports that are sent home. They are however kept

on file just in case someone makes a query. All this suggests a progressive dilution of the information that stems from classroom assessment. This starts building up when the teachers do not externalise all they learn from their assessment practices, and reaches the climax when just a percentage mark (i.e., the Assessment Mark) – without any teacher interpretation whatsoever (see Eggleston, 1991) – reaches the students' homes. Parents are consequently but passive recipients of minimal information sent by people outside the classroom (see Airasian; 2000; Filer & Pollard, 2000).

“

Michael's Story

Albeit I continually offer unofficial feedback to students (which is mostly verbal), the Assessment Marks are the only official evidence of classroom assessment to leave the classroom. Apart from the lack of space on the school's assessment report sheet, I see no point in writing comments that are neither taken seriously by the school administration nor sent to students' homes. Thus, unless parents contact me (which occurs very rarely, if at all), all they get to know, at least from me, is the Assessment Mark. When I began teaching at sixth form level, I decided not to initiate contact with parents for fear that students – who see themselves as 'mature' – might think I am 'spying on them'. This eventuality would, I believe, do more harm than good. I have reasons to believe that the parents may have the same misgivings about approaching me. For when a parent does get to contact me, I am more often than not asked not to say anything about this to the student concerned.

My experience suggests that students are generally more receptive to feedback when it is given on their own initiative. I still however try to give them as much as possible of the available information. I will use class tests as an example. Apart from the test mark, the corrected scripts normally contain the following: (i) the mark for each part question; (ii) brief remarks that are specific to particular 'situations' and which often invite the student to discuss certain 'difficulties' with me; and (iii) a concluding general comment regarding their overall performance. When I hand back the scripts, I normally dedicate the whole lesson to discuss the test. Following an overall comment about their performance, I go over the test paper and discuss briefly the individual questions – this is when I refer to the more common mistakes and/or the more interesting solutions that I would have noted. It is also at this point that I usually clarify the finer details of my marking scheme. In the final phase of this lesson, I usually give them some classwork, and as they work, I go quickly round the students to discuss their performance on a more personal basis.

”

7.6 The Beneficiaries and Benefits of Classroom Assessment

Other than the teacher and his or her students, the only other beneficiaries of classroom assessment mentioned by the participants were the school administrators and parents. Both the latter are however seen (and I cannot but agree) as recipients of

extremely limited benefits – administrators use Assessment Marks to determine in part the promotion of students, and parents learn about their children’s progress, even if only through the Assessment Marks. Although none of the participants referred to colleagues, some second year teachers (including myself) still manage to benefit from the assessment experience of colleagues when they access ‘past’ assessment records and information from the teacher previously in charge of their class (see section 7.2.1).

7.6.1 The Benefits to Teachers

The participants’ most frequently cited benefit of classroom assessment with regards to themselves was ‘the opportunity to learn about students’. Most of them maintained that classroom assessment renders them very knowledgeable about their students when it comes to ‘mathematics’. Their claims to knowledge extended to areas such as students’ understanding of mathematics, their mathematical preparation, their study efforts, as well as their mathematical potential and performance. It also appears that the more experienced the teacher is, the greater is the degree of self-confidence in his or her knowledge of students with regards to mathematics. Mario’s words portray the extreme confidence displayed by the more experienced participants:

By now [i.e., six months into the scholastic year], I have formed an opinion about each student in class. If you’ll point at any one of them, I’ll be able to tell you about his or her mathematical worth ... and my opinions are hardly ever wrong!

When Mario used the word ‘formed’, he was conveying the widely held notion amongst the teachers that they acquire this knowledge over a period of time that varies from a few weeks (e.g., Andrew insisted that after three weeks he would already know which students would fail the examination) to a few months (see also section 7.3.2). Like most of the teachers, Nicholas placed himself towards the less ‘speedy’ approach:

In the first month or two there were still some students that I didn’t really know ... but not now! I wouldn’t have done my duty as a teacher if I didn’t know all the students by now ... after teaching them for five months!

Nicholas’ notion that ‘it is the teacher’s duty to get to know his or her students mathematically’ – which echoes the creed of the ‘tried and tested practitioners’ that teachers should know their students inside out (see Gipps et al., 1995) – was a

recurring theme in every participant's discourse. By the time I conducted the 'assessment' interviews (i.e., between January and March of Year 3 of the study – see section 5.3.2), all the teachers asserted to know at least most of their students rather well as far as mathematics went. It was the 'middle ground' students about whom some teachers appeared to be less certain of their knowledge. Just like Weeden et al.'s (2002) 'invisible pupils', these are students with whom the teachers reportedly have few direct interactions, as they neither stand out in class by doing relatively well nor by being particularly weak. It was Rita who best brought out the participants' at times somewhat hesitant claims to knowledge with regards to these 'middle' students:

There are some students ... who are neither good nor weak, just somewhere in between ... about whom I feel less sure of myself. ... These students usually remain in the background ... they rarely open their mouths ... and I find little reason to approach them! What I know about them is based on what I observe, the homework (... but I'm never sure if they did it by themselves), and the test results ... but should they not come for tests, I'll really have big problems ...

To explore what the teachers actually 'know' about students, I invited them during the 'assessment' interview to pick two students from their class, and to recount what they know about them. The interview transcript below (in which Andrew talks about Jason, one of his students) illustrates the type of information that I came across:

Michael: ... *what can you tell me about Jason?*

Andrew: *He's quite good in mathematics ... he usually does his homework, pays attention in class, and gets good results in tests ... let me check [looks at his class records] ... just as I said, he got 75% in the first test, 70% in the second, and 85% in the third ... wish I had more students like him!*

Michael: *But ... what about his understanding of mathematics?*

Andrew: *Oh! I see what you mean ... I'm sure he'll pass the exam ... at the moment we're doing Trigonometry, and he seems to be following ... at least his work is usually correct, and he hardly has any serious difficulties ... Jason would let me know if he's not following ...*

Michael: *What can you tell me about his difficulties?*

Andrew: *You know! ... The usual things! ... A wrong sign here and there, fractions ... problems with Algebra ... but this concerns most of my students, not just Jason ...*

Michael: *Can you be any more specific than this?*

Andrew: *If I had some of his work in front of me, I would be able to pinpoint his mistakes ... but there's no way I could be more specific right here ... I never write these things down, and not just with Jason ...*

Michael: *What about his mathematical strengths?*

Andrew: *What can I say? ... he doesn't give up that easily ... and he seems to have quite a good mathematical background ... [refers to class records] ... he's got grade 2 in O-Level ... and that's not bad for an intermediate student! ... But I don't think I can be more specific than this!*

As it turned out, the teachers' knowledge of students appeared to be mostly related to their level of commitment to work, their behaviour in class, and their chances of examination success. Invariably, whenever I tried to 'press' for information that went beyond this 'ranking' to include evaluations against criteria (see Gipps, 1994), the teachers either ignored my efforts and kept to their chosen path, or else told me plainly that they could not provide such information. Another point that emerged from this exercise was that the teachers almost exclusively chose to speak about their better students (I ascertained this from the student assessment records that the teachers brought with them for the interviews). When I queried them about this choice (which reminded me of their declared preference to focus attention on the better students – see section 6.4), unless they attributed it to chance (as some did), the teachers maintained that these are probably the students about whom they are more knowledgeable.

A commonly claimed classroom assessment benefit amongst the teachers was that, given their student knowledge, they could fairly predict the examination results (i.e., the school's end-of-first-year examination [range of marks: from 0 to 70] and MATSEC's end-of-second-year examination [range of grades: from A to F]). To examine such claims, I asked the teachers towards the end of Year 2 of the study to predict a mark (first year teachers had to give a mark in multiples of 5 between 0 and 70) or a grade (second year teachers had to choose a grade from A, B, C, D, E and F) for each student, and to indicate in writing the criteria they used to arrive at these predictions. Table 7.9 gives for each teacher the computed correlation coefficient between his or her predictions and the students' subsequent examination results.

Table 7.9: The correlation coefficients based on teachers' predictions

First Year Classes		Second Year Classes	
Teacher	Correlation Coefficient	Teacher	Correlation Coefficient
Andrew	0.90	Carmel	0.68
Angelo	0.78	Kathleen	0.76
Jackie	0.81	Ray	0.73
Mario	0.71	Renzo	0.52
Matthew	0.78	Rita	0.58
Nicholas	0.87		
Stephen	0.81		
All Teachers	0.76	All Teachers	0.63

From what the teachers wrote, their predictions were largely based on the results of formal assessment (particularly, the class tests that reportedly mimic the very examinations whose results they were asked to predict – see section 7.4.2), even if the results of informal assessment were also taken into consideration (subsequent interview data revealed that the results of informal assessment were used here in pretty much the same manner in which they are being used to round off the Assessment Marks – see section 7.3.3). Table 7.9 reveals that whilst all the correlation coefficients are consistently positive and high (which in itself lends credibility to their claim of being able to predict examination results with a fair degree of accuracy), these tend to be generally higher for first year than for second year teachers (which, given the testing schedules identified in Table 7.2, supports the teachers’ notion that ‘the more one tests, the better the predictions’). Thus, whilst the literature remains undecided on the extent to which teachers can predict the examination performances of their students (see section 3.3.5), the present findings appear to add to the evidence in favour of those who argue that teachers can make good predictions. And they can apparently do so in spite of knowing so little about their students’ learning (see Weeden et al., 2002).

During the assessment interview I confronted the participants with their predictions and the actual examination results (the above correlation coefficients were never shown to the teachers). They appeared pleased but not surprised with the ‘overall closeness’ between the two sets of data. When they noted some strong ‘discrepancies’ between the predictions and the examination results, more often than not, they presented a series of plausible explanations that left unchallenged their predictions (e.g., Matthew ‘accused’ a student who performed much better in the examination [i.e., 43%] than he had expected [i.e., 5%] of copying). In fact, most of the teachers (especially, from amongst the more experienced ones) made it a point to maintain throughout the ‘validity’ of their own predictions over the examination results. This is how Mario expressed this idea:

I know what the students are capable of ... no examination result would ever convince me otherwise ... I'm 100% positive that my predictions are more truthful than the examination results!

By choosing to privilege their own predictions, these teachers confirmed once again the high esteem in which they hold classroom-based information (see section 7.2.2).

“

Michael's Story

Other than the students themselves (for I strongly believe that they have a deeper understanding of their own mathematical knowledge than they are normally attributed with), I find it highly unlikely that there should be anyone else who knows more about their mathematical knowledge than I do. But I still harbour no illusions of knowing all that much about them. I have reasons to believe instead that my supposedly 'privileged' knowledge of students is only very partial and shallow. Indeed, when I selected some of my students and tried to write down 'all that I know about them' with the help of my assessment records (along the lines of my request to the participants), what I could come up with were mostly generalisations. The very few instances in which I displayed some 'concrete' knowledge of students were largely based on the notes that I write after engaging in personal contact hours (see section 7.5.1) – but to which, however, only a few students normally attend.

Although I think that the ability to predict students' examination results is a positive thing in itself, it is not the 'deep' knowledge of students for which I yearn. In truth, this consideration has nothing to do with the fact that I registered the lowest correlation coefficient in the prediction exercise – a mere 0.37. I can think of two main reasons why my class test results (on which I, like my colleagues, largely based my predictions) are not very good predictors of the MATSEC examination result. First, I do not particularly plan my tests to mirror the MATSEC ones (even if I am lately starting to lay greater emphasis on MATSEC type questions in class – see section 7.4.2). Second, although experience has taught me that many students tend to prepare differently for class tests and examinations – in fact, they often admit later with me that they just barely studied, or not at all, for 'our test', but invariably add that this would not be the case with the MATSEC examination – I barely considered this in my predictions. The fact that second year class tests – unlike the first year ones on which promotion also depends (see section 2.4.4) – are low-stakes and thus unlikely to draw forth highly motivated best performances (see Torrance, 1995), may also help explain the lower predictive ability of us second year teachers (see Table 7.9).

”

7.6.2 The Benefits to Students

The participants concurred that classroom assessment could serve as an eye-opening experience for the students – an opportunity for them to face 'reality'. Most of the teachers added however that this whole exercise would be rather pointless unless it is followed (especially when the assessment feedback sends warning signals) by a strong resolve on the students' part to take note of the weaknesses identified, and to do something about them. Matthew was particularly keen to stress this point:

If a student is doing poorly, he ought to examine where his faults lie. He should ask himself, "Am I doing homework? Am I doing all of it ... or just the first question or two as I easily get fed up? Do I ask my difficulties in class? Do I

revise at home?” ... Things like that! The student won't be able to benefit from assessment unless he examines himself ... but then he also needs to act!

This 'acting on assessment' essentially means two things to the participants. First, that the students make a conscious effort to mend their ways (e.g., to start studying once 'lack of study' has been identified as a problem). Second, that the students keep them well informed about their difficulties (see section 7.2.2.3).

The teachers emphasised that the extent to which a student may actually benefit from classroom assessment depends solely on the student him or herself. They believe that sixth form students are generally too 'old' to be either effectively influenced or coerced in their actions by teachers, parents, or anyone else for that matter. Some teachers recounted personal experiences to show that when a student decides not to act upon the indications of assessment (as many of their students reportedly do), there is very little that they can do about it. Stephen provided one such incident:

I had a student last year who was frequently absent, hardly showed any interest when he turned up, never approached me with difficulties, and did very poorly in tests ... I kept warning him that he's bound to regret his behaviour (because I realised he had potential and ambitions) ... But only after he got his first 'disastrous' Assessment Mark did he decide to mend his ways ... He made lots of promises, and asked if he could come for 'contact hours' to make up for lost time. I agreed as I had completely believed his change of heart ... True to his word, he started coming regularly for lessons, paid attention in class, came prepared with difficulties during tutorials, and even came a couple of times for 'contact hours' ... But this change didn't last long! After a while he went back to his old self ... and then left school without even finishing the year.

This anecdote, apart from highlighting the widely held belief amongst the participants that many students apparently choose not to act upon the feedback received, also points towards another common notion amongst them – namely, that students, more often than not, tend to fall back on their own 'old ways' once their initial euphoria for change and improvement is over. Other forms of student reactions to negative feedback noted by the teachers – which essentially attest to the demotivating potential of assessment on students (see Broadfoot, 1996) – included: (i) labelling mathematics as 'a very difficult subject' and believing that it is beyond them; (ii) attending private lessons (and subsequently often losing interest in school lessons); (iii) changing the PMI option; and (iv) dropping out of school altogether.

The participants contended that all students, irrespective of how they react (or not) to its signals, stand to benefit from classroom assessment (albeit to varying degrees) both directly and indirectly. Starting with the **indirect benefits**, they argued that as teachers learn more about their teaching-learning situation through classroom assessment, they stand in an increasingly better position to gauge their teaching efforts to the fuller benefit of their students (see section 7.5.2). A number of teachers also claimed that some students ‘benefit’ indirectly when their parents ‘start forming an idea of what they are doing at school’ once the Assessment Marks start arriving home. These teachers opined however that, more often than not, the only manner in which the parents (perceived by all participants as ‘consumers’ of assessment who generally lack a sophisticated understanding of what is important in education – see Filer & Pollard, 2000) could ‘help’ is by encouraging students to attend for reparatory private lessons.

Although the teachers maintained that students stand to **benefit directly** from both formal and informal assessments (see introduction to section 7.3), they expressed a belief that students think to the contrary that they benefit far more from the formal than the informal. This belief – which is in line with Airasian’s (2000) observation that students tend to consider official classroom assessments to be very important – is reportedly based on what the teachers regularly observe in class. That is, their students’ much keener interest in formal assessment (primarily ‘class tests’) and its results. For instance, Mario contended that his students, at least those interested in their studies, expect and demand to be tested regularly, and are subsequently eager to learn about their test performance, especially the mark. The teachers attributed the student interest in class tests to two main reasons: (i) of all classroom assessments, tests carry the ‘heaviest consequences’ within the school system, as promotion from first year to second year also depends on them;¹¹ and (ii) class tests are the only form of classroom assessment that reflects the characteristics of the high-stakes external assessment (i.e., the end-of-first-year and the MATSEC examinations).

On their part, the teachers opined that students stand to benefit from class tests in three distinct ways (see Table 7.10) that all attest to their topmost ‘priority’ of helping students to pass examinations (see section 6.4). Albeit everyone mentioned these three

¹¹ The four teachers who do not publicise their Assessment Marks criteria in class (see section 7.1.2.2) opined that their students still come to realise that these marks are basically based on test results.

potential benefits, their greatest emphasis was invariably on the first one, which presents the capacity of class tests to motivate students as a means of individual control (see Broadfoot, 1996). The teachers' profound belief in the contribution of testing to students' learning (i.e., as represented by examination success) was again evident from the desire expressed by many of them for the school (or the mathematics department) to introduce half-yearly examinations for first year students and mock final examinations for second year students. The underlying idea was that as students would presumably take such 'semi-official' testing points more seriously than class tests, they are bound to 'learn' more and thus be better prepared for the examination.

Table 7.10: The potential benefits of class tests for students

- **Class tests keep students going:** Class tests, ideally spaced at regular intervals, are seen by the teachers as a means of keeping constant pressure on students in a bid to make them study throughout the whole year (as opposed to 'just a few days before the examination'). The teachers claimed that without class tests, students would slacken their efforts even more than is presently the case – something that, according to the teachers, would greatly jeopardise their examination chances.
- **Class tests show students where they stand:** Just as the teachers believe that all forms of classroom assessment help students to understand where they stand in the learning of mathematics and to plan ahead (should they so desire), they think that class tests are best suited to inform the students about their chances of examination success, as these two forms of assessment are closely related in both content (see section 7.4.2) and administration (see section 7.2.2.5).
- **Class tests prepare students for examination success:** Claiming that the process of preparing and sitting for a class test largely mirrors the examination one, the teachers opined that the more class tests students have, the better prepared they are for the examination. The most frequently mentioned benefits for students were learning 'to study for an examination' and 'to produce valid work under examination conditions'. In particular, Jackie gives tests that are more demanding than the examination itself in the belief that this increases the students' chances of examination success.

The link between student benefits and informal assessment appears to be far less linear in teachers' minds than it is for formal assessment. Indeed, some teachers – possibly because they are aware of keeping to themselves, particularly at the official level, a substantial part of what they learn about students (see section 7.5.3) – stressed that the measure in which students may learn about themselves from informal assessment largely depends on how much they are capable of capturing and interpreting its signals. Renzo made this point most forcefully:

If a student sets his mind to it, he doesn't really need anyone to tell him how he's doing ... all he needs is a 360° look around him! There's really no need to wait for the test results! If he follows lessons reasonably well, answers to my questions, does his classwork and homework without major problems, and can work the past papers on his own ... then he should know that he's doing OK! The student has to figure this out for himself ... I cannot do it for him!

Renzo's comment was one of the few instances in the study when the participants somewhat referred to the notion of student self-assessment – their understanding being that students may 'well gauge where they stand' should they reflect upon their own 'performance' in assessment activities, particularly the informal ones. Even though the teachers disclosed an impression that most probably only few students actually go through this introspective process, especially on a systematic basis, none of them reported to have ever tried to help students develop this capacity (see Stefani, 1998).

“

Michael's Story

I agree with my colleagues that students stand to benefit from classroom assessment both directly and indirectly. I also subscribe to some of their views, namely that: (i) students need to act on their assessment feedback; (ii) students take a keener interest in formal assessment than in informal assessment; and (iii) class tests may serve students along the lines detailed in Table 7.10. But then, I do not concur with their position that the extent to which a student may benefit from classroom assessment depends almost exclusively on him or herself. I believe instead that the teacher should never shift his or her assessment responsibilities onto students. Thus, even though I too would like them to keep me constantly informed about their difficulties, I still try to actively learn about their strengths and weaknesses. Albeit I feel the need to help students learn how to monitor their own learning, my efforts in this direction are as yet sporadic and normally only involve the few students who come for personal contact hours. This is when I have more time to discuss self-assessment and how to go about it. The closest I come to engaging my students in peer assessment is during some classroom discussion, especially when we 'reconvene' to discuss the work of particular groups on nonroutine tasks (see section 7.2.2.4).

”

7.7 Constraints and Explanations

7.7.1 Teachers on the Quality of their Assessment Practices

The participants displayed very little of the uneasiness about their assessment practices that, according to Stiggins (1992), is a characteristic of many teachers. With the

exception of Kathleen (who argued that she still has much to learn) and Carmel (who specifically criticised his questioning techniques), all the others spoke rather positively and confidently about their classroom assessment practices, conveying in the process an aura of competence and self-approval. To justify their self-confidence, many teachers referred to their alleged ‘deep’ mathematical knowledge of students as allegedly demonstrated by their proven ability to predict the examination results (see section 7.6.1). The teachers’ main argument was that their assessment practices are sufficiently good to provide them with more than adequate information about students’ learning and the effectiveness of their teaching. Andrew subscribed to this position:

Modesty apart, I’m quite pleased with what I do ... I’m not saying there’s no room for improvement ... there always is! But my assessment practices help me to get to know the students well and inform me about my teaching ...

Albeit the comments about their information generating practices (see section 7.2.2) were generally positive, the teachers sounded possibly even more positive about their practices that serve purposes other than that of generating information (e.g., grading, responding to assessment, and communicating results). Even Kathleen sounded much surer of herself when she spoke about her non-‘information generating’ practices:

What can I say? ... I’m just doing what needs to be done! ... I see to their difficulties, I compute the Assessment Marks and pass them on to the administration ... I speak to students who need to be warned ... things like that! ... I just don’t see how I can do things differently or better!

Even though most teachers claimed to be largely satisfied with their practices, all of them frequently referred to a number of ‘circumstances’, ‘conditions’ or ‘constraints’ that affect negatively at least some of their practices. The teachers’ awareness that pressures within a given social context have a mediating effect on assessment practices (see Broadfoot, 1996) does not however dampen their rather positive outlook. It generates instead a ‘proud feeling’ in many of them – a sense that they are ‘doing a very good job in spite of the difficult circumstances’. Andrew insisted on this:

When I say I’m pleased, you’ll have to keep in mind that I’m working under certain circumstances ... My point is that in spite of these circumstances or constraints ... call them what you like ... I’m still doing a good job! But obviously, I would do it even better should I be working in better conditions!

Most teachers likewise shared the belief that they would improve on their already good assessment practices should the constraints that presently impinge upon them either become less felt or cease to exist. This prevailing ‘feel good’ factor may explain their general reluctance to enrol in assessment training programmes (see section 6.3.1).

“

Michael’s Story

Hard as it is to admit, I am not at all pleased with my assessment practices, and this irrespective of how they compare to those of my colleagues. My research endeavours, which served to make me reflect upon my own practices, have led me to this ‘negative’ self-awareness – an unanticipated development that I have often described as ‘upsetting’ in my fieldwork journal. With hindsight I realise that it could not have possibly been otherwise, as I was now adding to my teacher lens the even more discerning lens of the researcher to review my practices. This ‘unseating conflict’ between my ‘teacher’ and ‘researcher’ personae helped me however eventually to empathise more with the participants, and in the process to gain deeper understandings into their practices and explanations.

”

7.7.2 The Constraints on TCAP

A number of factors were found to govern the participants’ selection of classroom assessment tasks (see section 7.4.2). These factors were classified under national, school, and personal contexts (see Table 7.5). But whilst the teachers tended to refer to the factors from the national and school contexts as ‘constraints’ – in the sense that they “delimit a teacher’s field of activity” (Woods, 1990, p. 15) – this was not the case with the factors from the personal context. Some teachers moreover even spoke of the ‘personal’ factors that guide their task selection with a great sense of pride. A case in point was the self-assured manner in which Angelo admitted to exclude Tasks 3 and 4 because of the unorthodox and ‘non-mathematical’ style (of which he clearly does not approve) in which they are written:

I don’t accept them ... I’m not accustomed to such things! Give me rather something like ‘solve the equation’ or ‘form an equation’ ... not these trivialities!

The participants’ tendency to see constraints ‘only in things distant from themselves’ emerged again when we discussed classroom assessment beyond task selection. Apart

from Kathleen who suggested that her ‘limited’ assessment knowledge probably constrains her practices, the constraints identified by the other teachers hardly ever had to do with them personally, even if only remotely. Nicholas’ suggestion that he may ‘lack the time and initiative’ to try out new things was one of the few occasions when the perceived constraints were not rooted in the ‘national’ and ‘school’ contexts. The teacher-identified constraints listed in Table 7.11 are somehow related to the classroom assessment practices that reportedly inform the teachers about their teaching-learning situation. This suggests that from all the various facets of classroom

Table 7.11: Constraints on classroom assessment

- **Large student numbers:** All the teachers view the ‘large’ number of students in PMI classes (which can be up to 50) as a barrier to getting to know students mathematically. Their main concern is that with such numbers they cannot give individual attention to students in class, especially when it comes to observing them work and to addressing their difficulties.
- **Inadequate O-Level preparation:** The teachers agree that secondary education is not preparing students sufficiently well to further their studies in mathematics, even if only at I-Level. They thus find it difficult – also because many students, apart from being unmotivated, are weak in mathematics and English – to present students with ‘challenging’ tasks (see section 7.4.2).
- **Lack of personal time:** Some teachers described themselves as being very busy people outside school hours (see section 6.2). As a result, they claimed not to have enough time to prepare new materials (e.g., worksheets and tests) and to mark student work at home.
- **Lack of contact time:** The teachers concurred that syllabus pressures do not permit them to engage in classroom activities (no matter how beneficial they judge them to be) that require an extensive amount of classroom time. Some teachers even explained their declared focus on the better students in terms of time constraints (see section 6.4).
- **Lack of space:** The teachers complained that classrooms are generally too small for the number of students in PMI classes (see section 2.5.1). They alleged that this condition constrains their assessment practices in two ways: (i) finding it difficult to move freely about the classroom to either observe students at work (see section 7.2.2.1) or to check their work (see section 7.2.2.2); and (ii) making it relatively easy for students to copy during class tests.
- **The Matriculation Certificate:** The teachers agreed that students are generally overburdened with the six subjects of the Matriculation Certificate (see section 6.3.5). They argued that this ‘burden’ makes it hard for students to do all (or even most of) the set work, and that this failure subsequently triggers off a series of assessment reactions – namely, that students are less likely to ‘discover’ their difficulties and to make them known to the teacher, who finds him or herself in turn in a less favourable position to respond in an appropriate manner. Most teachers also complained that students hardly have the time at school to come for personal contact hours.
- **The examination:** Although all the teachers referred to the examination questions (particularly the MATSEC ones) as a constraint on their choice of assessment tasks, only three of them (i.e., Andrew, Nicholas and Rita) reportedly view the existence *per se* of the examination as an assessment constraint. They spoke of the examination as an ‘unfortunate reality’ with which they however have to live.

assessment, the teachers judge themselves to be primarily constrained when it comes to matters that have to do with either accessing or generating assessment information.

“

Michael's Story

My experience suggests that the various constraints identified by my colleagues do in fact have a negative impact on one's assessment practices. But then, I do not agree with their commonly held notions that assessment constraints originate almost exclusively from within the national and school contexts, and that these constraints affect mostly those practices that inform teachers about their teaching-learning situation. To start with, I hold that one's personal context exerts a notable influence on one's practices – the first example that comes to mind, however trivial it may seem, is the debilitating effect over the past few years that my PhD work has had on the energy and time that I could afford on my school work (classroom assessment included). Besides this, I am well aware that the harmful effects of the constraints on my assessment practices go well beyond jeopardising the quality of assessment information that reaches me – it is something instead that permeates throughout all that I do (or fail to do!). .

”

7.7.3 Explanations for TCAP

Anderson (2003) contends that teachers base their decisions on ‘what they have always done’ and on ‘real and practical constraints’. True to this, the participants essentially explained their assessment practices at these two different, albeit constantly interactive, levels – namely, ‘what they believe in and know about assessment’ (which recalls Anderson's first reason) and ‘what circumstances permit’ (which is Anderson's second reason). Matthew was very clear about this:

Michael, you must understand that my practices do not necessarily go hand in hand with what I want or like ... in life one has to compromise! In truth, what I do in class is mostly what I want ... but there are things that I would rather do but stop from doing because of circumstances ...

The teachers reportedly view this process as one that is primarily based on the former (i.e., their knowledge and beliefs), but then only to the extent that the latter (i.e., the circumstances) permits. This balancing act between ‘ideals’ and ‘reality’ shows how the participants' assessment decisions are greatly influenced by their deep ‘sense of practicality’ (see Hargreaves, 1994) as reflected in their habit of ‘considering the

consequences’ (see Anderson, 2003) before acting. Even teachers, like Jackie, who came out as very resolute and protective with regards to the ‘worth’ of their assessment practices, repeatedly conceded that an array of circumstances do not permit them to do all they believe in. What Jackie said about her tests was a case in point:

I’m a strong believer in very demanding tests ... I would otherwise be misleading my students in believing they can make it, when it’s not the case! Tests should prepare them for the hard examination reality! ... However, as a concession to weak students, I usually start the test with a couple of relatively easy questions ... they’ll have no excuse then for getting a zero! ... That’s the only reason I do it!

In Kathleen’s case, this ongoing interactive process between ‘ideals’ and ‘reality’ is reportedly less based on the former than is generally the case with the other participants. This probably results from her strong belief that she is ‘still at the initial learning stage’ with regards to assessment matters. As a matter of fact, Kathleen is the only teacher to see training as a truly valid means for improving her assessment practices (see section 6.3.1). All the other teachers are of the opinion that such improvement is barred by the incumbent circumstances or constraints (which they see as almost exclusively external to them – see section 7.7.2). They hold that it is only after the removal or the toning down of these constraints that one can actually envisage progress in classroom assessment. Matthew had this to say:

If I had less students in class, if they were better prepared, more intelligent ... and if ... I can keep going because my list of complaints is long! I can tell you that if these things weren’t there, then my assessment would be better than at present!

“

Michael’s Story

Like my colleagues, I find that my assessment practices basically embody what remains of my knowledge and beliefs after these are sifted by the circumstances (or constraints) in which I work. Albeit I agree with them that constraints impede the quality and development of TCAP, I do not share their almost unanimous stance that improvement in the quality of these practices depends almost solely on improvements in the embedding national and school contexts. This line of argumentation is greatly reductive. For whilst contexts external to the teacher certainly exert a considerable constraining influence on TCAP, the impact of the teacher’s personal context should neither be neglected nor underestimated.

”

7.7.4 TCAP in a Constraint Free Environment

After discussing with the participants the manner in which they feel constrained in their classroom assessment practices, I invited them to project themselves into an 'ideal situation' wherein all their previously identified assessment constraints would not exist. This exercise meant to gain an understanding into the type of assessment practices that the teachers would be most likely to adopt in an environment that they deem to be constraint-free. What came out most clearly from their responses was that in such an eventuality each one would practically retain all of his or her current assessment practices. Angelo claimed this by saying:

I think I would still be doing the same things that I'm doing now ... I may do something a bit more or a bit less, but ... I can't imagine doing anything that is really different from what I'm doing now ...

These declared intentions are further evidence of how much the teachers are rather convinced of what they are presently doing (see section 7.7.1). Seen from a different perspective, one can say using Anderson's (2003) phraseology that, contrary to what their previous comments on the current situation appeared to suggest, their 'real and practical constraints' do not seem to constitute that much of an influence on 'what they have always done'.

But most teachers, just like Angelo, still maintained that in an ideal situation they would most probably affect some modifications to their practices. The first thing to emerge was that the modifications in teachers' minds are solely restricted to those practices that reportedly inform them about their teaching-learning situation. In addition, the bulk of these would-be modifications have to do with the 'frequency' or the 'quantity' of the targeted practices – for example, Stephen would give less class tests, Andrew and Nicholas would dedicate more time to classroom discussion, Rita would give more worksheets, Kathleen would check students' work more frequently and thoroughly, Ray would ask more questions in class and set more work, and Carmel would work more questions in class. To a much lesser extent, the teachers also mentioned modifications that are related to the format (e.g., Jackie would introduce multiple choice questions) and the 'level of complexity' involved (e.g., Mario would include harder class test questions). True to Stiggins and Bridgeford's (1985) claim

that few teachers are prepared to explore and use new assessment approaches, although most of the participants had previously claimed that they would not use Tasks 3 (or similar ones) primarily in view of the existing constraints (see Table 7.5), it was only Andrew who actually mentioned its use (albeit still only occasionally and solely for the purpose of classroom discussion) in a constraint-free environment.

All this suggests that the teachers' would-be classroom assessment modifications are meant, more than anything else, to consolidate their already existing practices. In reality, the teachers emerged as being primarily interested in doing 'better' the things they are already doing. A number of participants did however refer to some 'new' assessment practices (which basically relate to how teachers obtain information about their teaching-learning situation – see Table 7.12) that allegedly cannot be used at

Table 7.12: The 'new' TCAP in a constraint free environment

- **Oral mathematics:** Andrew, Jackie and Nicholas spoke of introducing an oral component to classroom assessment. The emphasis here would be that the teacher listens to individual students as they 'think aloud' about how they would go about solving problems. As Jackie explained, the whole purpose would be "*to get a glimpse at students' mathematical reasoning, how they tackle a question ... something that often eludes me when I look at their written solutions*".
- **Student presentations:** Andrew and Stephen suggested that small groups of students could be assigned some research task (e.g., going to the library and learning about 'simple Matrix operations') to be performed under the guidance of the teacher. Their idea was that these students would eventually present their research results (which could take the form of 'proving something' or 'working some unseen questions on the whiteboard') to the other students. The teacher would then assess and grade their presentation. Implied in this approach is that students work together in groups (a possible assessment development to which Nicholas also referred).
- **Projects:** Renzo opined that students could be asked to work on a project (either individually or in a small group) that would ideally be linked to their future university studies. For instance, he suggested that should a student intend to study a commercial subject at university, he or she might be asked to write a long essay on 'The use of calculus in commercial subjects'. Renzo argued that this, apart from helping the students to appreciate more the relevance of mathematics, would provide the teacher with useful insights on his or her students' wider understandings of mathematical applications.

present in view of the existing constraints. But more than actually expressing an intention to use these new practices should the constraining circumstances change,

these teachers spoke vaguely instead about the possibility of introducing them. All the indications thus remain that the classroom assessment situation would not change drastically in a constraint-free environment in spite of the teachers' oft-repeated claim that this would. On the contrary, given the teachers' declared propensity to confirm their present practices with but a sprinkle of modifications, the resulting assessment scenario is very much likely to be more of the same as at present with just the possibility of the odd 'novelty' on the side.

“

Michael's Story

If I had to project myself just a few years back (i.e., before I started to deepen my understanding of assessment issues in general, and classroom assessment issues in particular), I think that I would have responded to the request that I made to the participants in a manner that would have been very similar to theirs. For I, like them, used to think that the impeding constraints on my assessment practices lie 'outside myself' – a notion that easily leads to a debilitating 'there's nothing I can do about it' kind of attitude. Now I know differently. Whilst I do not deny that the 'outside' constraints exist, I find that the 'inside' ones (i.e., those that originate from someone's personal experiences, beliefs and views) have an equally, if possibly not more, incapacitating effect over classroom assessment practices. If anything, the constraints that originate from within are easier to overlook (being often so integral to a person's mental framework) and consequently harder to address.

”

Part FIVE

DISCUSSION AND CONCLUSIONS

CHAPTER 8

Exploring the Implications of TCAP for Learning

8.0 Getting Focussed

As from now, I will no longer keep bracketing ‘my story’ away from those of the twelve ‘official participants’. Instead, I will weave ‘mine’ into ‘theirs’ in a manner that highlights the similarities and contrasts amongst our stories – a process that further enlightens my understanding of TCAP. The main research findings related to TCAP presented in Part Four may be classified under eight, closely interrelated categories. It can be said that, generally speaking, the assessment practices used by the participants:

- lack an effective educational base;
- pursue traditional forms and approaches to teaching;
- emphasise non-professional functions;
- are firmly in teachers’ hands;
- are characterised by collegial isolation;
- are not well integrated with teaching and learning;
- provide teachers with surface knowledge of students; and
- provide students and outsiders with limited feedback.

In the coming section, I will briefly revisit TCAP in relation to these eight encompassing categories. This exercise serves to refocus the main research findings in such a way as to facilitate the exploration of the implications of TCAP for learning.

8.1 TCAP – Revisiting the Main Findings

8.1.1 Assessment that Lacks an Effective Educational Base

Although most of my participating colleagues have completed some form of initial teacher training and/or attended the occasional short in-service course, the vast majority maintained that they either had no training whatsoever in assessment or else that they could not remember if they had. In a scenario that largely mirrors the lack of assessment literacy amongst teachers reported in previous studies (see section 3.3.2), the three teachers who mentioned some form of assessment training recounted that this simply consisted of a one-hour lecture. In my case, had it not been for my present

studies, I would probably still lack an assessment theoretical base. None of my colleagues attributed the origins of any of their assessment practices to formal training. Instead, they explained these practices through a variety of experiences at both personal (first as a student and later as a teacher) and communal (i.e., the cyclic hand-over of assessment practices from one generation of teachers to the next) levels. Whilst my assessment practices have also been fashioned by similar experiences, I am currently striving to translate my gained expertise into improved practices. These, however, remain as yet not that much different from those of my colleagues.

My colleagues' insistence on the impact of experiential knowledge on the quality and development of their assessment practices – which lends credibility to the frequently made claim (see Cizek et al., 1995) that much of what teachers know about assessment is learned 'on-the-job' or through 'trial-and-error' – reflects their widespread belief that experience in teaching is the 'real key' to professional success. Consequently, it is experience in assessment rather than training that these teachers seek and value. Besides their general unwillingness to explore and use new assessment approaches – a teacher reluctance that also emerged in a study by Stiggins and Bridgeford (1985) – their assessment practices lack a direct reference to theoretical underpinnings, particularly those of the emerging assessment paradigm. Their practices (including Ray's experiments) are based instead on what they judge to work for them inside their own classrooms – a consequence of what Hargreaves (1994) calls the 'sense of practicality amongst teachers'. In fact, any changes in assessment practices (the essence of which has been found not to change once these become established) are more or less pragmatic reactions by my colleagues to ongoing developments in their classroom experiences. There is indeed no evidence to suggest that any such changes reflect a desire on these teachers' part to transform their assessment practices.

8.1.2 Assessment that Pursues Traditional Forms and Approaches to Teaching

To learn about our teaching-learning situation, we basically rely on fleeting and distant observations, on very rare monitoring of students' work, on largely closed questions, on students' initiative to make their difficulties known, and on class tests consisting of 'routine' tasks that truly constitute a central feature of our classroom assessment practices. These tasks – contrary to being 'significant' (see NCTM, 1995) or

'authentic' (see Eisner, 1993; Murphy, 1996) – tend to be close-ended, highly scaffolded, lower order, topic specific, time efficient, not related to real life situations, and contain minimal language. Other than my occasional forays into 'open' assessment situations, nonroutine collaborative tasks are systematically excluded from inside classrooms. Moreover, if and when my colleagues' students work together, both inside and outside the classroom, it is on the students' own initiative – something that the teachers tolerate, if at all, rather than encourage. As with informal assessment, marked work that is not clearly the effort of individual students is not given any real weighting when we report at an official level, such as in the computation of the Assessment Marks (which are consequently largely based on class test marks).

Our official assessment results thus perpetuate the psychometric notion that student performance 'measured' under strict examination conditions produces a mark that, save a few exceptions, accurately reflects what students know. The widespread importance that my colleagues attach to testing in assessment is such that whilst they cannot imagine classroom assessment without tests, some of them find it difficult to talk about their assessment practices without continually referring to class tests. Much along the lines of 'convergent assessment' (Torrance & Pryor, 1998) and 'evidence gatherers' (Gipps et al., 1995), these teachers use in turn what they learn from assessment to sustain and promote their traditional views of teaching as transmission and learning as practice. In contrast to my budding efforts aimed at using assessment to promote a constructivist learning environment, my colleagues use assessment evidence to check if their 'message' is getting through. Just as Black (1998) alleges, when this is judged to be the case, at least with the better students on whom they reportedly count, these teachers then move on to present new material. Otherwise, they keep repeating (and/or rephrasing) the message and give additional practice work until such time when the students who count (practically those deemed to have a talent worth investing in – see Torrance, 1995) have mastered it.

8.1.3 Assessment that Emphasises Non-Professional Functions

Our assessment practices appear to be greatly influenced by accountability and managerial concerns. Much along the lines reported by Mavrommatis (1997), my colleagues tend to see assessment largely in terms of selection, certification and

accountability, rather than in terms of its professional 'for learning' role. Instead of seeking to establish a 'workable relationship' between their summative and formative roles (see Black, 1998) – something on which I am working – these teachers usually favour the former over the latter. Moreover, as they feel accountable to a number of people – including students, colleagues, school administrators, parents, and the general public – they try to deal with classroom assessment in a manner that asserts publicly and unassailably their professional competence and commitment. They thus choose assessment practices that maximise students' chances of examination success (which they perceive as their ultimate teaching goal and what everyone expects from them) and discard others that do not, even when these are judged to be beneficial for learning. In contrast, my late attention to MATSEC-type questions, particularly towards the end of the year, followed the realisation that my students were paying the price for my 'inattention' to examinations. But then, like my colleagues, I also produce 'dead data' for reporting purposes (see Broadfoot, 1996) that are seen to be fair and are publicly defensible. It is moreover largely due to accountability concerns that we all hold back from making official use of assessment knowledge that cannot be easily defended on psychometric grounds (see Gipps, 1994). Albeit I add what 'I know' about a student whenever an official request for information is made, this does not alter the fact that my official results, like those of my colleagues, are mostly based on what a student is able to achieve at a particular point in time (namely, performance in class tests), rather than on what or how much a student has learned over time.

My colleagues and I also use assessment as a means to manage and control students. Assessment indeed helps us to reward or admonish students, to make them study, to select them for second year studies, and to prepare them for the examination. A few even mentioned using assessment to move forward the lesson. It thus appears that assessment serves us to 'manoeuvre successfully' our classroom operations amongst the various pressing accountability and managerial demands at class, school and national levels. My colleagues believe that by so doing they are not only upholding the expectations of all concerned, but are also protecting their own interests and reputation from a possible barrage of accusations linked to favouritism, unfair practices, and professional incompetence. With survival within the system being so important for teachers in general (Cole, 1997; Day, 1993), it appears that classroom assessment offers these teachers an unsuspecting means for achieving this. For apart from possibly

achieving teaching satisfaction through it (see Airasian, 2000), it offers them an opportunity to demonstrate ‘proven’ impartiality and professional competence whilst pursuing what can be described as widely accepted and much cherished goals. All this is linked to the issue of ‘sameness’ in assessment – my colleagues practically use the same assessment practices with all students largely out of concern for fairness, particularly from the point of view of ‘others’. The fact that I still do not organise my assessment practices according to the students’ individual needs, at least not as much as I would like, has however mostly to do with time constraints.

8.1.4 Assessment that is Firmly in Teachers’ Hands

Albeit we feel answerable to various people – basically the very ‘clients’ identified by Calderhead (1987) – my colleagues consider classroom assessment as something that concerns primarily and benefits directly just the teacher and students (both of which should not, in my opinion, be the case). In spite of acknowledging the students’ direct role and interest in classroom assessment, we still largely position students at the receiving end and us in almost complete control of this process. We practically decide the ‘why’, the ‘what’, the ‘how’, and the ‘when’ to assess, as well as how to interpret and respond to the results of assessment. Our students, who are never consulted and not even always informed about the desirable performance standards or the assessment criteria being adopted, are expected instead to follow closely, almost blindly, the assessment lead indicated by the teacher. In line with this strong student assessment reliance on teachers (see Sadler, 1989), although my colleagues view self-assessment (or a semblance of it) as a positive quality possessed by a few students, they neither recognise it officially nor do they actively strive for its development. They also appear to be completely unfamiliar with the notion of peer assessment. On my part, although I encourage students to make their own sense of the classroom assessment process, I only sporadically discuss with them how they can go about this metacognitive process, and I do not give student self- and peer assessment any official recognition in class.

With regards to classroom assessment and ‘classroom outsiders’, my colleagues show little affinity with the many calls in the literature (e.g., Black, 1998; Elwood & Klenowski, 2002) for the creation and development of ‘communities of practice’. They tend to relegate the role of these persons to a decidedly distant, non-influential

one – being there just to offer support when asked to. Whilst all of them are determined not to lose out on their much enjoyed and cherished classroom autonomy (see Hargreaves, 1994), which extends beyond assessment to include most of their actions inside the classroom, the only assessment-related difficulty that some of them see in it concerns the subjectivity involved in awarding the Assessment Marks. However, although some colleagues are clearly concerned that their assessment ‘freedom’ may jeopardise the fairness of the school promotion system, they go no further than to call for school or departmental guidelines or criteria that govern the award of these marks. Their willingness to somewhat curtail their assessment autonomy when classroom assessment carries direct, tangible repercussions outside the classroom – which largely echoes Gipps’ (1994) and Harlen’s (1994a) ‘reliability’ concerns and recommendations – does not however alter the overall picture characterised by each teacher regarding assessment as his or her personal prerogative (see Sadler, 1989). Although I do not assent to this prerogative, no classroom outsider has any control whatsoever over my classroom assessment practices.

8.1.5 Assessment that is Characterised by Collegial Isolation

Speaking of teaching in general, my colleagues claim to be almost completely in the dark about what the other teachers do inside their own classrooms. This strong measure of ‘autonomous isolation’ (see Gipps et al., 1995) – which emerges as one of the more pronounced teacher characteristics within the department – appears to permeate all spheres of our classroom practices, assessment included. Within this scenario, teacher assessment ‘collaboration’ is mostly restricted to situations in which a novice teacher seeks the advice of his or her more experienced colleagues until the day when he or she decides to have gained enough personal experience to be able to continue on his or her own. Apart from these early attempts at acquiring ‘recipe knowledge’ in a bid to move from novice to expert (see Hargreaves, 1993), there are otherwise only flashes of assessment collaboration amongst us (e.g., teachers sharing assessment information or discussing their class tests and results) that are mostly restricted to persons within the same office. More often than not, we start each new scholastic year without any kind of assessment information, official or otherwise, about our students, and subsequently keep to ourselves most of what we learn about the students throughout the year.

With little or no assessment dialogue amongst us, some of my colleagues claim to get a glimpse at other teachers' classroom assessment practices through past or present mutual students. Although possessing, at best, very limited knowledge of colleagues' assessment practices, some teachers are still highly critical of these practices, particularly with regards to the setting of class tests and the computation of the Assessment Marks. This form of collegial 'mistrust' contrasts in turn with the rather positive manner in which my colleagues generally speak about their own practices. There is also some evidence that the fear of being negatively judged by one's colleagues is channelling some of them to do (or avoid) certain assessment practices that they would otherwise avoid (or do). It thus appears that these teachers fear for themselves the same potential harm of labelling that is normally linked with students following traditional assessments (see Gipps & Stobart, 1993). Rather than collaborate on assessment, my colleagues appear on the defensive, as if they are in competition with one another. They may however be surprised to learn that, albeit we operate in undeniable isolation that helps to engender mistrust amongst colleagues (see Cole, 1997), our overall classroom assessment scenarios hardly differ.

8.1.6 Assessment that is Not Well Integrated with Teaching and Learning

Like teachers in general (see Calfee & Masuda, 1997), my colleagues are clearly more interested in teaching and content coverage than assessment. In line with the transmission approach to teaching and learning (see von Glasersfeld, 1989), teaching is for them a continual dishing of information by the teacher for students to capture, and learning is gauged by students' ability to memorise and reproduce this information on demand. Assessment, under its various forms, is basically then a means for these teachers to establish the extent to which, if at all, their teaching endeavours are having the desired effects on students – basically what Brown (1989) describes as an 'effectiveness check' on the communication from teacher to students. In reality, these teachers view teaching, learning and assessment as three distinct phases that follow one another in a largely linear, disjointed fashion. They view this process as beginning with teaching, then students either learn or fail to learn, and only subsequently does assessment determine the success or failure of teaching and learning. In this essentially behaviouristic approach to educational practice (see Shepard, 1991), assessment appears to stand outside teaching and learning, instead of in dynamic interaction with

them as part of a flowing, mutually illuminating process – the very notion on which the new assessment paradigm is built (see Gipps, 1994). Although I hold constructivist beliefs that presuppose close integration between instruction and assessment, my inability as yet to move away from a teaching style that can be largely categorised into the separate phases of ‘exposition’, ‘practice’ and ‘consolidation’ (see section 6.5.3) renders this integration inside my classroom still somewhat problematic.

My colleagues’ ‘objectified’ sense of assessment (see Ellis, 2001) is such that even though they claim to find assessment, particularly the informal, an indispensable classroom tool that permits them to gain insights into their teaching-learning situation, assessment emerges as an event or object that stands apart from teaching and learning. Not only is the influence of assessment on teaching and learning limited, but it also appears to be largely restricted to situations involving their better students. In fact, apart from the occasional claim that they cannot possibly do anything about what they have learned from assessment, their actual responses to assessment are restricted to consolidation that consists primarily of repetition and practice work, encouragement or warnings to students, finding the right teaching level (as different from ‘style’), and slight changes in assessment practices. Moreover, whilst these teachers view assessment as an opportunity for students to understand (generally, on their own) where they stand and to plan ahead, they never use assessment to seriously question their own classroom practices. This treatment of own practices as unquestioned and unquestionable (see Schön, 1983) strongly suggests that assessment is for these teachers primarily about and for students, not themselves. This much-challenged ‘standing outside the assessment discourse’ (see Torrance, 2000) permits them in turn to act as disinterested judges of current student achievement and future potential.

8.1.7 Assessment that Provides Teachers with Surface Knowledge of Students

Almost all my colleagues, particularly the more experienced ones, maintain that assessment provides them with deep knowledge of students as far as mathematics is concerned – claims that are generally built on the premise that their sustained interaction with students guarantees such knowledge. The proven ability, of some more than others, to largely predict the examination results is also frequently cited by many of them as tangible proof of their deep knowledge of students. Still, some

teachers readily admit that they feel much less confident about their knowledge of certain students, usually the ones who do not stand out in class as doing relatively well or particularly weak. But other than this, it can generally be said that – contrary to Airasian's (2000) advice – my colleagues have a lot of faith in their own assessments. Albeit I do not hold to know my students any less than my colleagues do (and this in spite of my lesser predictive ability), I am still far from satisfied with my student knowledge. My experience suggests that it is the quality of interaction between teacher and students, not interaction *per se*, that is important. Indeed, I gain the deeper knowledge when I meet students individually or in very small groups during personal contact hours. But since very few students actually use contact hours, and only very occasionally, my knowledge of students is more likely to be related to what they can retain and reproduce from lessons than to what they know and can do in mathematics.

In spite of my colleagues' claims, the research evidence – in line with Black's (1998) observation that teachers do not usually have sound information about students' progress – suggests that their knowledge of students, even of those they claim to know well, lacks depth. Their knowledge appears, at best, to be mostly related to students' level of commitment to work, their class behaviour, and their chances of examination success. It can hardly be otherwise given their inclination to assess recognition or recall of facts rather than level and complexity of understanding. My colleagues moreover tend to generalise their understandings from particular assessments, which often enough only target a narrow sample of students and a narrow sample of each student's behaviour, to other forms of assessments and much wider contexts. Their tendency to speak in general terms rather than in terms of the specific assessment set – which is quite common amongst teachers (Airasian, 2000) – probably explains why they often speak of students' competence rather than performance. Whilst there is no evidence to suggest that these teachers use assessment to help students produce their best performance, they insist with students to 'reveal themselves to them' and bitterly complain that this happens only by a few students and not often enough. Just like my colleagues, my knowledge of students is also hindered by an assessment record system that much relies, similar to that of the 'intuitives' (see Gipps et al., 1995), on the teacher's memory of what students can do. Two of us however keep short written records of some of our encounters with students – a practice that, albeit short of Murphy and Torrance's (1988) formative recording, partially addresses this situation.

8.1.8 Assessment that Provides Students and Outsiders with Limited Feedback

Although assessment is an ongoing inside our classrooms – we assess informally on a continuous basis during the normal course of lessons and formally on more specific and far in between occasions – the assessment ‘feedback’ that students and, to a much greater extent, classroom outsiders receive appears to be limited. Moreover, it is feedback that often serves the social and managerial functions at the expense of the learning function (see Weeden et al., 2002). For it seems to lack the gap-closing qualities that give classroom assessment its formative dimension (see Ramaprasad, 1983; Sadler, 1989). In fact, the only assessment communication outside the classroom, at least officially, consists of grades (i.e., Assessment Marks) without any teacher interpretation. Students are however exposed as insiders to wider and deeper feedback than outsiders. For sure, there is what they manage to learn by themselves – even if the chances are that they are neither trained nor guided by their teacher to conduct self- or peer assessment. The feedback that students receive from informal assessment situations – especially during lessons, as it is normally richer during tutorials and personal contact hours – consists mostly of teachers nodding or shaking their heads, passing encouraging remarks or rebukes, and clarifying difficulties. Over and above this, feedback from formal assessment carries marks. Particularly in view of my colleagues’ overriding notion that it is then up to each individual student to draw his or her own conclusions and to act upon them, it may even be that we are primarily offering ‘monitoring’ rather than ‘feedback’ (see Wiliam & Black, 1996). All this suggests that students (and outsiders even more) come into contact with only a small portion of the feedback possibilities that classroom assessment has to offer.

Our feedback appears to be at its ‘weakest’ with the weaker students. With a number of colleagues making it a priority to interact more and to give more attention to their better students (i.e., those who presumably stand a better chance to pass the examination), they are likely to get to know less about their weaker students and also to have less opportunity to pass on feedback to them. This has parallels inside my classroom. For it is with the same few students that approach me individually at my office, who are usually the more motivated as opposed to being the better ones, that I produce the richer feedback. All in all, although my colleagues appear to be very appreciative for all they learn from assessment, be it through informal or formal

practices, and would like students to keep them well informed about their difficulties, they do not seem inclined to share the full extent of this acquired information with their students, or with anyone else. This is particularly the case with what they learn from their informal everyday assessments – which are basically ‘qualitative judgements’ (see Sadler, 1989). These teachers indeed show restraint in communicating feedback that is not backed by marks, as they think that such feedback is hard to defend and is also taken less seriously by those on the receiving end. Their implied equation between grading and good, publicly defensible assessment – a teacher characteristic also noted by Airasian (2000) – probably also affects negatively the quality of the feedback they offer. For with informal assessment constituting the huge bulk of classroom assessment activity, this ‘holding back’ suggests that they may be keeping to themselves much of their assessment-based professional knowledge. This works against presenting feedback that can be understood and used by students – what Stiggins and Conklin (1992) call ‘high-communication value’ feedback.

8.2 The Impact of Teacher Beliefs on TCAP

My experience does not support Huberman and Miles’ (1984; cited in Hargreaves, 1994) assertion that practice changes before beliefs. For I have so far largely failed to translate my newly acquired assessment knowledge and strong personal convictions (which conform to the new educational vision being proposed for Malta – see section 2.6.2) into standard classroom practices. Indeed, I still only venture but tentatively and very occasionally into the less traditional forms of assessment that parallel my constructivist beliefs and my desire for reformed educational practices (as opposed to ‘just new policies’ – see section 9.0) in Malta. Needless to say, I find this conflict between my beliefs and actions highly disturbing. Contrary to my resulting dissatisfaction, my colleagues largely welcome and often take pride in their assessment practices, which are generally even more traditional than mine. The evidence suggests that my colleagues’ primarily positive outlook on their own assessment practices is linked to their ‘privileged’ position of basically assessing in harmony with their beliefs about the educational process in general and mathematics in particular. In the next two sections, I will explore the impact of these essentially very traditional beliefs on TCAP.

8.2.1 Teacher Beliefs about the Educational Process and TCAP

The educational discourse of my colleagues reveals a very traditional understanding of the principles that govern and inspire the educational process. These teachers – irrespective of age, teacher training, teaching experience and all other differentials – emerge as true and convinced representatives of the educational thought that regulates the procedures and practices of the behaviouristic philosophy that still dominates the local educational scene (see section 2.6.1). The ongoing emergence of new policy documents that seek to align local education with progressive practices, whilst attesting to the disenchantment at official level with the current dominant philosophy, do not yet appear to have won the hearts and minds of Maltese teachers, certainly not of my colleagues. To the contrary, my colleagues appear opposed to initiatives that steer away from their deeply ingrained behaviouristic beliefs that developed as they ‘succeeded’ through the system first as students and later as teachers. Indeed, they regard the introduction of the differentiated papers at O-Level (see section 2.2.1.1), the matriculation system that brought curricular breadth at sixth form level (see section 2.2.1.2), and the increasing number of students proceeding to pre-tertiary studies (see section 2.1.2) as threats to quality and traditional standards rather than as the intended opportunities for learning. At the root of their discontent lies the notion that education is a process that gradually sifts and rewards people on the basis of ‘ability’ and ‘willingness to work hard’ – two qualities they claim to possess. This explains why the slow but constant democratisation of the local educational process goes against their understanding that, as ‘ultimate success’ is not even within the reach of most students, it makes no sense to keep encouraging more and more students to continue studying beyond the mandatory school age in sixth form colleges (see section 6.3.3).

Sensing that policy makers, both at school and national levels, are doing little to safeguard the traditional high standards that allegedly accompanied their scholastic journeys, these teachers apparently see in classroom assessment an opportunity to somewhat redress this situation from within. Their insistence is thus on ‘salvaging what they can’. Rather than seeing assessment as the means to promote learning for all, they choose to narrow their assessment attention on the ‘better’ students – basically those who according to their calculations stand a reasonable chance of eventually being ‘successful’ in the system. For the rest of their students, assessment

becomes an event that primarily signals failure and criticism, if not a not so subtle invitation to reconsider the continuation of one's studies. My colleagues thus seem to react to the undesired increase in student numbers by mentally, if not physically, eliminating those – not few in reality – that in their opinion should not be there in the first place. This 'sidelining' – which occurs in spite of many colleagues reportedly feeling sorry for the weaker students, especially the ones who make an effort – appears to be doubly justified in their eyes in view of the links between our school and University (see section 2.4.1). For these teachers take the school's affiliation with University to mean 'elitism'. This, however, is not the case. The University, instead, became involved primarily to ensure a planned and gradual transition between secondary schooling and tertiary studies (see Buhagiar, 2003). Consequently, the school's link with University – contrary to what my colleagues appear to think – has nothing to do with restricting the student numbers to include only the better qualified.

8.2.2 Teacher Beliefs about Mathematics and TCAP

The general sidelining of the 'weaker' students in class is undoubtedly also linked to the 'elitist' manner in which my colleagues view mathematics. They consider mathematics, especially at post-secondary level, as an abstract, highly structured and theoretical subject that is consequently only suitable for intelligent, rational, motivated and hard-working students. For them, mathematics is certainly not meant for all students (see Gates, 2002). It follows that they consider the efforts directed at teaching the 'unteachables' as a waste of time, both for the teacher and the students themselves. With such students, assessment assumes the role of helping students 'understand' this reality – in particular, that although the system has allowed them (against the teachers' better advice) to pursue the study of mathematics beyond O-Level, 'facts' are now showing the futility of this exercise. This reasoning is built on their premise that assessment-based evaluations, especially if derived from tests, are highly likely to be precise and true. In turn, their image of mathematics as a deeply hierarchical subject practically condemns to 'beyond redemption' those students who repeatedly do not achieve. The understanding is that although the teacher still learns through assessment about their weaknesses, there is little that he or she can do at this point to rectify the situation. In line with their behaviourist beliefs, my colleagues think that these students have too many missing 'building blocks' in their learning history to be able

now to move successfully to higher levels (see Gipps, 1994; Shepard, 1991). It is as if such students – who, like the rest, are seen by these teachers as empty receptacles to be filled – have reached a point at which the highly respected ‘teach-learn-assess’ cycle (see section 8.1.2) can no longer help them to learn.

This sense of helplessness does not however lead my colleagues to reconsider their behaviourist teaching and assessment methods with which they grew up and made it through the system. Their own successful stories – all claimed to have performed consistently very well in mathematics at school – appear instead to give these methods an aura of efficacy and approval. Consequent to this ‘guarantee’, my colleagues tend to attribute failure in mathematics to either the student being weak, unprepared or unmotivated – or a combination of the three. But I know from personal experience that this ‘distancing by teachers’ is grossly unfair with students. Unlike my colleagues, there have been times when I experienced failure in mathematics (see section 1.0). I know with hindsight that this failure arose primarily from my resistance to link doing mathematics simply with either rote learning or following procedures correctly. In fact, when I chanced on a teacher who appreciated that the development of meaning or understanding is a personal process that depends on what the student makes of successive experiences (see Watson, 2002), I began to realise that even I can do mathematics, and rather well for that matter. My struggling periods – which were invariably accompanied by assessment episodes that inhibited rather than contributed to my learning – encourage me now as a teacher to persevere with students in difficulty. I know that their ‘failure to learn’ probably reflects my inability to provide them with experiences that respect their unique ways of learning (see Gardner, 1991 [cited in Gipps, 1996]; LaCelle-Peterson, 2000). My colleagues – for whom doing mathematics has always been a relatively ‘simple’ thing – however appear to lack this type of sensitivity. They believe in one way to learn mathematics – the way that ‘worked’ for them. These teachers moreover try to achieve this learning by inundating students with practice tasks that are primarily relevant for the examination, not for life (see section 7.4.3). These tasks, apart from largely failing to stimulate the analytical, critical and creative thinking skills of students, do not help them to develop their mathematical understandings into knowledge that can be used in real-life contexts (see Murphy, 1996). It is thus not only the ‘weaker’ students who are missing out on a quality education, but also those who supposedly win within the system.

8.3 The Implications of TCAP for Learning

Having established the main characteristics of TCAP and recognised the impact of traditional teacher beliefs upon them, I now move on to explore the implications for learning of these practices. The aim is to examine TCAP in relation to the learning needs of the students currently involved. Because if assessment is to be ‘for learning’, it must operate formatively for both the teacher and his or her present students (see section 3.2.3). The implications formulated here subsequently serve to shed light on the quality of TCAP from the perspective of the emerging assessment paradigm.

8.3.1 When Assessment Lacks an Effective Educational Base

When pre-service professional courses give assessment only a cursory mention at best, novice teachers can only draw on their own experiences and those of their colleagues – usually the senior ones who may themselves have little, no or outdated assessment training. Whilst teacher collaboration on assessment can lead to the beneficial ‘guild knowledge’ of which Sadler (1989) speaks, a mostly cyclic inter-teacher dependence along the lines identified in this study is possibly counter-productive because it may limit assessment knowledge and practices to those of current (Bright & Joyner, 1998) or past teachers. This almost complete reliance on accumulated craft knowledge leads to an assessment scenario dominated by traditional practices. Not only does this not encourage teachers to explore and use new assessment approaches, but it increases the chances that they emphasise practices that are not particularly conducive to learning (e.g., reliance on closed questioning) and exclude others that are (e.g., open questioning). Such a reality is mostly detrimental to students. For it affects the content they learn (i.e., emphasis on examination type material), the way they learn it (i.e., through transmission followed by practice), and the quality of judgments made about their achievement (i.e., largely summative) (Schafer, 1993).

The system’s failure to help teachers become assessment literate makes it impossible for them to get the full benefit of classroom assessment (Gipps, 1994). Teachers cannot assess well without understanding the constructs that they are assessing, without knowing how to get at student’s knowledge and understanding, and without knowing how to elicit the student’s best performance (Gipps, 1994). Failing this,

teachers are likely to continue assessing to assign grades rather than to learn about students – what Bright and Joyner (1998) term ‘assess to assess’. This constitutes a great threat to their ability to draw valid inferences about students’ learning that are needed to address the day-to-day decisions inside classrooms. Indeed, when teachers make the measurable important instead of making the important measurable, they are likely to come in contact simply with the easily measurable as opposed to desirable, valid information (Murphy & Torrance, 1988; Wiliam, 2001).

8.3.2 When Assessment Pursues Traditional Forms and Approaches to Teaching

The traditional forms of TCAP identified here favour an atomised approach to learning (see Black, 1998). This approach, which in Marton and Säljö’s (1976) terminology is ‘surface’ as opposed to ‘deep’, accentuates learning by rote, of small pieces of information without the understanding that interrelates them, and of fixed rules and procedures that are grasped only as tactics, without the strategic overview needed to give them significance and to guide their application (see Murphy, 1996). When students, as in this study, are constantly exposed to decontextualised, rote-oriented tasks that impose low cognitive demands rather than meaningful learning (Darling-Hammond, 1994), the development of their conceptual knowledge suffers (Boaler, 1998). For sure, this reliance on ‘recall’ tasks – as the use of practically the same (or very similar) closed tasks for teaching, practising and assessment purposes renders even the most complex task a matter of recall (see Brookhart, 1999) – makes it unlikely for students to engage in Vygotsky’s Zone of Proximal Development (ZPD). This is a zone within which judiciously scaffolded activities, such as tasks that are neither too trivial nor too demanding, make it possible for learners to reach higher competencies through personal constructions. But teachers’ failure to present meaningful and purposeful classroom activities works against students finding, creating and negotiating their own meanings (Murphy, 1996).

The traditional TCAP depicted here lack the multiplicity and variety that are needed to focus on the full range of achievement targets (i.e., knowledge, thinking, processes, products, and dispositions) and to reflect the understanding that different students show what they know and can do in different ways. These practices consequently

cannot provide the deep insights into students' understanding that have the potential to stimulate further learning (see Torrance & Pryor, 1998). And as assessments assiduously target individuals, students are not being given the opportunity to involve themselves in collaborative projects that create the conditions for thinking aloud and sharing ideas, which is an important metacognitive aspect of learning (Ellis, 2001). Again, when teachers, as appears to be the case in this study, fail to give sufficiently challenging work, students start producing performances well below their real competencies in the belief that they cannot do more, that they have reached their limit (Torrance & Pryor, 1998). But this is unlikely to be noted and/or cause concern. Because even though students do not reach the intended learning goals (e.g., making decisions, planning routes through tasks, choosing methods, and applying mathematical knowledge – see Boaler, 1998), they still develop problem-solving strategies that help them pass examinations (Brookhart, 1999).

8.3.3 When Assessment Emphasises Non-Professional Functions

When, as in this study, TCAP are primarily motivated by the maximisation of students' examination success, assessment focuses on students' performance rather than understanding. The ensuing prominence to marks and grades, which emphasises the individual's achievement in relation to others (usually without defining what a student knows and can do) possibly leaves learning to what von Glasersfeld (1995) calls 'fortunate accidents'. For sure, the accompanying 'teach to the test' phenomenon basically just promotes the learning of skills and competencies that can be incorporated within the traditional paper-and-pencil examination questions. Given the 'learning limitations' of the test-like tasks that are normally used in class (see section 8.3.2), it appears that students are not being helped to develop the 'deep competence' of which Schoenfeld (1999) speaks. In a situation characterised by students "not so much described by tests as constructed by them" (Hanson, 2000, p. 68), students are encouraged instead to think that it is only worth learning those things assessed by testing (see Bryant & Driscoll, 1998). In view of the decontextualised, rote-oriented tasks that traditionally dominate the testing scenario, students may thus be misled into thinking that mathematics is simply about learning procedures and descriptions as opposed to also learning about connections and relationships with previous knowledge (see Marton & Säljö, 1997; cited in Watson, 2002).

The summative slant that the teachers studied generally give to most of their assessments inhibits the freedom and attention to individual needs that formative work requires (Black, 1998). By seeking sameness in assessment – which allegedly gives them an aura of fairness – these teachers miss out on differentiated practices that would serve to identify and to help meet the learning needs of each student (see Perrenoud, 1998). When assessment is mostly presented to students as an instrument of selection, it is easier to foster in them rote learning, cheating, and competitive and grade-hunting attitudes than an appreciation that assessment, including self- and peer assessment, can also facilitate learning. The teachers' failure to balance their formative and summative roles – apart from possibly straining their relationship with students who may see them more as judges than facilitators (Gipps, 1994) – can also serve to estrange students from schooling (see ARG, 2002; Broadfoot, 1996). The use of assessment as an instrument to manage (or even coerce) students and their learning in turn defeats the understanding that learning flourishes within an equitable environment that is characterised by individualised assessment dialogue between teacher and student. Learning is also better served when – contrary to the 'reputation related holding back by teachers' so evident in this study – all available information is used for action that best meets the particular learning needs of each student (Black, 1998).

8.3.4 When Assessment is Firmly in Teachers' Hands

By not opening the classroom assessment process to public scrutiny and modification, teachers miss out on an opportunity to enhance assessment, and consequently learning (see NCTM, 1995). For assessment standards are bound to improve when all stakeholders are allowed to participate in assessment development (Phye, 1997b), as opposed to only 'letting them in' on the very rare occasions when circumstances dictate this. But even more damaging for learning is the teacher practice of almost invariably relating students to the receiving end of assessment. When student involvement in assessment is simply limited to 'providing the data' for teachers – with no further involvement either prior or after this phase – it is most unlikely for effective learning to take place. For learning falters when students simply rely on receiving feedback from teachers, even if this has good formative potential (Sadler, 1989). Effective learning not only requires learners to have some part in deciding goals and identifying criteria for assessment progress in order to understand what it is that they

are trying to achieve and to want to achieve it, but also to subsequently engage in peer and self-assessment (ARG, 2002). Failing this level of ‘agency’ on the part of students (see Murphy, 1996), understanding suffers because when students do not monitor and reflect on their own work/performance, it is unlikely that they become self-monitoring learners in the metacognitive mode. Other ‘victims’ of this assessment reality are student commitment and motivation (ARG, 2002; Brookhart, 1999). With demotivating assessment effectively pushing students out of the system (Broadfoot, 1996), it seems that learning is being penalised both directly and indirectly when teachers do not allow students to share in the responsibilities of assessment.

This exclusion of students also carries what is possibly an even more damaging long-term consequence for learning. By failing to develop the skills of self-assessment in students, teachers miss out on an opportunity to equip students with the desire and the capacity to become self-directed lifelong learners. Of its very nature, self-assessment engages students in self-reflection and in the identification of the next steps in their learning – which is, after all, how independent learners seek out and gain new skills, new knowledge and new understandings (ARG, 2002). Even the very notion of formative assessment requires students to assume a central role in their assessment, as this helps them to understand their strengths and weaknesses, and how to deal with them (Harlen & James, 1997). The capacity of students to judge their own work – particularly in view of its propensity to develop complex understandings through reflective habits of mind – is consequently more than a bonus in good formative assessment (see Black, 1998). Indeed, deep learning develops when students engage in thinking about the meaning of what is being learned – a metacognitive process that helps them to take control over their own learning (Gipps, 1994).

8.3.5 When Assessment is Characterised by Collegial Isolation

When teachers choose, as in this study, to maintain professional isolation rather than engage in communities of practice, this carries repercussions – firstly for themselves and consequently for learning. With regards to teachers, an isolationist mentality is likely to undermine the development of their reflective dimension. Communities of practice, instead, engage teachers in reflective practice that serves to critically examine and analyse their beliefs and practices for clarification and resolution of meaning

(Elwood & Klenowski, 2002). In particular, a mathematics classroom assessment process that is open to participation by colleagues would help teachers to form common definitions of the mathematics to be assessed and to reach consensus on the appropriate evidence of students' learning (NCTM, 1995). Such collegial involvement in assessment helps the individual teacher to develop the skills to analyse and evaluate students' work, as 'challenges' from other teachers force him or her to be specific about what students' misconceptions are and to articulate standards for students' work (Bryant & Driscoll, 1998). And as teachers become more assessment literate, the more likely that their assessment practices are to be for learning (see section 8.3.1).

Collegial assessment collaboration can also give rise to several new approaches that are not even anticipated in the literature (see Black et al., 2003). As teachers stimulate and challenge one another, as they share ideas that worked and others that did not, they come to experience what Brookhart (1999) defines as 'some of the most interesting and useful information'. Collaboration – which does not imply that teachers adopt and adapt the same practices – helps to transform each teacher's individual accumulation of assessment knowledge into 'guild knowledge' (see Sadler, 1989). Teachers can thus build up mutual expertise that no one else possesses, and that no one else can give them (Black, 1998). But by failing to engage collegial communication, as is largely the case in this study, each teacher – however experienced – remains solely in possession of his or her conceptions of quality assessment and knowledge of current students. In this scenario, assessment practices are probably transmitted from one generation of teachers to another, from one teacher to another, without much effort to engage in either personal or group reconstructions. TCAP thus become 'fossilised' – ever less likely to change in favour of new ones that are more conducive to learning.

8.3.6 When Assessment is Not Well Integrated with Teaching and Learning

When assessment, as often emerges in this study, stands outside teaching and learning, it is unlikely to meet what NCTM (1995) sets as the primary goal of classroom assessment – that is, to advance students' learning and to inform teachers as they make instructional decisions. For sure, there is little opportunity for the teacher and students to interact and arrive at shared meanings along the lines of the 'systematic planners' (see section 3.3.3.2). Assessment that is not aligned to instruction works against the

need for instructional decisions to be based on information collected before, during, and after instruction (see Anderson, 2003). It follows that assessment needs to be part of effective planning, with the teacher having a clear vision of his or her learning goals, and how these can be translated into assessments that make sense. In an integrated approach, assessment has to be flexible to respond to initial and arising instructional needs, and teacher planning should include strategies to ensure that learners not only understand the goals they are pursuing and the criteria that will be applied in assessing their work, but also receive feedback, take part in assessing their learning, and are helped to make further progress (ARG, 2002). This approach examines how well 'all' students are doing, what targets they should set to improve learning, and what strategies might get them where they want to go.

On the contrary, an assessment that does not feed into planning remains a detached activity devoid of opportunities for both learners and teacher to obtain and use information about progress towards learning goals. There is moreover less possibility that improvements in assessment lead to improvements in the process of teaching and learning at classroom level (Torrance, 1995). Again, the focus of assessment is likely to be on its product, its results, and the manner in which these results may be used to manage or even drive school systems rather than on the learning process itself and how this can be facilitated (see Torrance, 1995). This defeats the constructivist purpose of having assessment and instruction in dynamic interaction (see Gipps, 1994). For learning does not progress through the simple, isolated act of measuring, but through the complex network of decisions and actions that links assessment, learning and motivation (Weeden et al., 2002). In particular, assessment is for learning when the information it produces is used to address the learning needs of the current students (Black et al., 2003). This builds on the understanding that when assessment information is recorded or used for long-term curriculum improvement, it may be formative for the teacher, but not for his or her present students.

8.3.7 When Assessment Leads to Surface Knowledge of Students

The assessment efforts of the teachers studied are unlikely to produce information about what students actually know and can do, let alone about what they can nearly do (see Gipps, 1994). Given that good instructional decisions depend to a large degree on

the quality of the evidence collected (Airasian, 2000; NCTM, 1995), these teachers do not appear well placed to use assessment formatively, even if they so wish. For sure, they lack the necessary valid and detailed information, especially about those students whom they admit to know very little. Teachers need instead to examine and reflect on 'rich' evidence in order to determine whether or not there is a gap between actual and desired levels of performance, and which actions can successfully close this gap (see Wiliam & Black, 1996). To do otherwise, teachers run the risk of ending up with information that offers too little and too late to be of any real formative use. Limited knowledge of students moreover works against the teacher offering students appropriate scaffolding activities, which are neither too trivial nor too demanding, that provide students with the needed support and guidance to perform at a higher level (Murphy, 1996). This would help students to close the gap between what they can do on their own and what they can do with the help of others.

For teachers to decide successfully what to do next, which students to push, and so on, it is thus not enough that they try to make sense of the available evidence that is derived from assessments that either 'just happen' or else from assessments that either do not have the potential to produce rich information (see Wiliam & Black, 1996) or whose potential is not exploited enough. In reality, whilst 'meagre' evidence may well serve a number of non-professional purposes (the participants' ability to use test results to fairly predict examination success is a case in point), it is not suitable for instructional purposes (see Brookhart, 1999). To be for learning, assessments need to promote valid inferences about learning through adequate and relevant evidence (NCTM, 1995). In other words, no amount of reflective examination by the teacher can ever turn extremely poor quality evidence into information that helps to further learning. Such evidence can however lead to what Bright and Joyner (1998) call 'illusions of learning'. These are the over-generalisations made by teachers of the depth of understanding that is actually present in what students say and do that arise from teachers' often strong desire for students to understand.

8.3.8 When Assessment Provides Limited Feedback

With teachers, as in this study, assuming almost total responsibility for classroom assessment, student learning becomes even more dependent on teacher feedback. This

makes it even more important that feedback is intended and has the potential to improve learning. But when teacher feedback, as is presently the case, is largely limited to grades or short normative comments (usually of a simple praise or blame nature) without any reference to mastery and progress, it can only offer at best a most rudimentary indication of where one stands in comparison to others or norms. The understanding that feedback ought to help students close the gap between where they actually stand and where it is desirable that they be (see Ramaprasad, 1983; Sadler; 1989; also Black et al., 2003) means that, in most cases, the manner in which the teachers studied respond to what they learn from assessment does not even qualify as feedback, but is more akin to monitoring (see Wiliam & Black, 1996). It is simply not enough for assessment to just indicate the existence of a gap. This would deny students from what should be for them the more fundamental aspect of assessment – that is, to understand through user-friendly information how they are doing and how specifically they might be able to improve upon what they are doing (Stefani, 1998).

Moreover, should teacher feedback be in terms of what has not been learned – as is normally the case when assessment is concerned with finding out what facts and skills have been acquired rather than understanding (Harlen & James, 1997) – the chances are that it may even serve to demotivate students. This is likely to happen when students with low self-esteem perceive such information, which can never be seen as a means of closing the gap, as yet another confirmation of their inability to perform. Bearing in mind that when assessment demotivates students it effectively pushes them out of the system (Broadfoot, 1996), it may well be that the ‘not for closing the gap’ feedback acts as the ultimate barrier to students’ learning. The learning situation is further aggravated by the barest possible assessment information that is generally made available to classroom outsiders, parents included. As these persons lack well-documented assessment information (which is characterised by ‘focus’, ‘friendliness’ and ‘forthrightness’ – see Anderson, 2003), they cannot enter into an effective relationship with teachers in support of students’ learning (see Weeden et al., 2002).

8.4 The Quality of Teachers’ Classroom Assessment Practices

Given the crucial link between classroom assessment, teaching and learning (see ARG, 1999), quality in assessment is increasingly being defined in terms of how much

it actually assists the learning process (Broadfoot, 1996). It is from this perspective that I now examine the quality of TCAP. But before presenting what the research findings actually suggest, I first revisit briefly what the teachers in the study say about the quality of their practises. .

8.4.1 What the Teachers Say

In contrast to my continuing misgivings about my own assessments, my colleagues seem to be engaged in assessment practices that largely meet their ideals. They generally see themselves as competent classroom assessors and see little need, if at all, for any assessment training or retraining. These teachers moreover tend to consider themselves as singularly knowledgeable about their teaching-learning situation in view of the sound knowledge of students, as far as mathematics goes, that they reportedly derive from classroom assessment. Their claims to knowledge are mostly based on the belief that they are in the best position to collect extensive information about students over an extended period of time. Often enough, they back their claims to knowledge of students, and hence to the good quality of their assessment practices, by parallel claims of an ability on their part to predict the students' examination results – an ability that a correlational analysis in this study has somewhat confirmed.

My colleagues' largely positive outlook on their assessment practices is reflected in the deep reliance on their part on the indications that these provide. Indeed, these teachers overwhelmingly maintain to be far more convinced of and ready to rely on assessment evidence generated and recorded within their own classrooms than anything else, including the results of the school and certification examinations. This is in line with their firm conviction that they are doing a good assessment job given the circumstances, some of which they see as constraining (see section 7.7.2), in which they operate. The propensity of my colleagues to value the quality of their current practices is also evident from their declarations of how they would like to assess in an ideal, constraint free environment. For their main aspiration is to do better the things they are already doing. Their primary interest in consolidating the existing assessment practices strongly suggests that these teachers have conservation, rather than change, on their minds (see Hargreaves, 1994).

8.4.2 What the Research Findings Suggest

Although the classroom offers much promise for formative assessment (Nuttall, 1993), this study corroborates Black's (1998) observation that the potential of classroom assessment to improve teaching and learning is still not being exploited. Not only do the teachers studied appear to be largely unaware of the limitations of the various methods they use (as with all other data collection methods for that matter) to collect information about their teaching-learning situation, but their chorus of well-established practices also emphasise 'assessment of learning' with a clear intent on measuring, grading and reporting that has little to do with the use of assessment for the promotion of learning (see ARG, 1999). Indeed, the TCAP presented here more than parallels the 'legacy of the past' that continues to drive the current assessment situation in US schools (see Delandshere, 2001).

This study indicates an assessment reality that remains alien to the emerging view of assessment that is tentatively encapsulated in alternatives such as 'educational assessment' (see Gipps, 1994) and 'assessment for learning' (see ARG, 1999, 2002; Black et al., 2003; Weeden et al., 2002). Instead, it is by far closer to the traditional notion that knowledge is static, universal and monolithic (see Delandshere, 2001). These are the very philosophical foundations upon which 'convergent assessment' (see Torrance & Pryor, 1998) and the 'intuitives model of the tried and tested practitioners' (see Gipps et al., 1995) – the two teacher assessment models that have close affinities with the TCAP presented here – are based. As emerged in this study, the classroom assessment discourse within such a scenario is primarily concerned with 'passing and failing', 'knowing and not knowing', and defining who becomes one sort of person and who another. This emphasis on measurement and comparison occurs within the wider recognition that educational assessment is the unquestioned arbitrator of value, whether of students' achievements, institutional quality or national educational competitiveness (Broadfoot, 2000). Assessment assumes within this perspective the role of reproducing social stability by distributing social and economic goods in a fashion that is only supposedly meritocratic (Torrance, 2000).

From an assessment for learning perspective, the evidence strongly suggests that the TCAP presented here contain a number of characteristics that undermine the quality of

the whole process (see section 8.3). The research findings thus indicate a diametrically opposed evaluation of the quality of TCAP to the one depicted by my colleagues (see section 8.4.1). The most likely explanation for this discrepancy is that these teachers hold a different conception of quality in assessment to the one being championed by the proponents of the emerging alternative assessment paradigm. For they appear to link quality assessment mostly with adequate monitoring of and timely support to the notion of teaching as transmission and learning as practice, and the measuring of what students 'know' under the guise of a mark that is a good predictor of examination performance. This conception contrasts sharply with the present understanding that quality classroom assessment is essentially one that can be used formatively with and by current students, makes reference to important performance or learning criteria (with the 'self' being the main criterion), and takes place in authentic settings. Such an assessment process is moreover led by the teacher with the direct participation of his or her students – an opportunity for the teacher to cash on assessment that supports the teaching-learning process, and for students to reflect on their learning and to gain a measure of ownership over their own assessment.

8.5 Recognising the Need for Change

Sections 8.4.1 and 8.4.2 present two clearly contrasting evaluations of the quality of the same TCAP identified in this study. These strikingly different interpretations of practically the same evidence, albeit captured from decidedly different angulations, cannot but reflect the different positions or values held by my colleagues (i.e., the participants) and myself respectively (see Sowder, 1998). These two sets of values basically mirror the notions of 'assessment of learning' in my colleagues' case and 'assessment for learning' in my case. As a practising teacher myself, I do appreciate and understand what motivates these teachers to speak in the way they do. For I also understand the importance of experience and the pressures of moving the lesson forward so that the syllabus can be finished on time, of preparing students for the examination, of appearing to be fair to everyone, of producing defensible grades that the school administration (as well as others) expects, of monitoring the teaching-learning situation and intervening when appropriate – this is a list that can go on and on. But, at the same time, as someone who has deepened my knowledge in assessment and its intimate links with teaching and learning, I also recognise the hefty price on

learning that the prevailing assessment situation within our department is extracting. In particular, the negative implications for learning presented in section 8.3 encourage me to work in favour of directing the classroom assessment discourse towards practices that serve primarily as a platform for action in support of learning.

This is not a denial on my part that teachers, at least for the foreseeable future, also have a summative role to perform inside classrooms (see Black, 1998). I am guided instead by an understanding that, as Gipps and Murphy (1994) argue, the main purpose of assessment should be to enhance the educational process (i.e., professional) rather than to manage the educational system. From this perspective, it can be argued that the research findings indicate a strong need for change. At stake are necessary and important instructional decisions that need to be based on good evidence (Airasian, 2000; Brookhart, 1999). Failing this, invalid conclusions about the success of instruction may be drawn with harmful consequences for students (Airasian, 2000). Given that teachers are the key to educational change (Hargreaves, 1994), the fact that my colleagues tend to consider their assessment practices in a rather positive light and show little intention of affecting any real changes to them may well constitute a major problem. For as Fullan (2001) maintains, if change is to be implemented successfully, it is important that the people involved in the change process perceive or feel the need for change. But then, even though teachers are undoubtedly classroom frontlines, they are by no means the only hurdles that need to be addressed in order to bring about change in classroom assessment. My own experience has been rather revealing in this respect. Although there were times when I even ‘accused’ myself of not having tried well enough, of having been too busy with this study to concentrate on transforming my assessment practices, I still do not think that my lack of significant progress can be adequately explained, at least for the most part, through lack of time (see section 9.2.2). My experience suggests instead that “we must [also] change existing conditions so that it is normal and possible for a majority of people to move forward” (Fullan, 2001, p. 269). The question thus remains to determine what needs to be done to move TCAP closer to assessment for learning within the type of difficulties encountered and in the knowledge that “there can never be a recipe for change” (Fullan, 2004, p. 54). This means that the proposals for improvement that I put forward in the next chapter consciously draw on what ‘works’ *in situ* rather than the ‘nonexistent’ absolute truth.

CHAPTER 9

Moving towards Assessment for Learning

9.0 The Reform Process in Malta: What do the Research Findings Suggest?

Chapter 8 highlighted the need – at least, from the value position that I have come to privilege – for assessment change in the teachers studied. Much along the lines reported in other countries (see Black & Wiliam, 1998; Delandshere, 2001), the TCAP presented here appear to be largely ‘untouched’ by the numerous calls to bring classroom assessment primarily at the service of learning. Most of these teachers appear moreover to be opposed to and to resist the tangible manifestations that the current reform efforts in Malta – which seek to provide a quality education for all and to widen the access to post-compulsory studies (see section 2.6.2) – have so far put across their path (see section 2.6.3). This emerges from their non-acceptance of the differentiated paper system at SEC level, their general reluctance towards the PMI option, and their wide concern about the increasing number of students at pre-university level. It is as if these developments, which these teachers invariably link to lowering of standards, go against what they see as ‘normal’. Namely, that as students move up the educational ladder, the ever fewer successful some will proceed to higher studies, whilst those who fail will either continue with a diluted study programme or else, if they have already reached employment age, seek a job that matches their lack of higher qualifications. But this reality, characterised by the misguided notion that ‘assessment of learning’ renders it meritocratic (see section 2.6.1; also Torrance, 2000), is exactly what the Maltese policy makers have been trying to steer the local educational system away from for the past ten years or so.

This study consequently suggests that in spite of the many published policy documents, media communications and ensuing legislation, the much publicised reform process has probably still not reached where it matters most – that is, the teacher/classroom level. For not only have the classroom practices of the teachers studied remained practically unchanged, but these teachers also appear to be largely unfamiliar with and/or unmoved by the spirit of the intended reforms. By itself, this finding sheds important insights on the current implementation strategy of the local

reform process in which ‘assessment for learning’ features prominently. It emerges clearly that improved policies, even if backed by legislation, do not necessarily lead to improved practices (see section 3.3.3.3). There is not even the guarantee that teachers actually get to know the details of the new policies, let alone understand and desire them. The top-down local practice of having a few ‘experts’ drawing a policy document on the initiative of the Minister of Education and only subsequently organising meetings, amongst others, with teachers to primarily convince them of its worth (the latest example is the school networking exercise that is currently being proposed – see Ministry of Education, Youth and Employment, 2005) does not appear to be bearing fruit. In reality, one reform seems to follow another with very little to show at classroom level. It may even well be that when teachers, as in this study, largely fail to appreciate the intended reform vision and make it their own, they are very likely to see any proposed or imposed changes as attempts by classroom outsiders to score either political or administrative points rather than as genuine efforts aimed at improving their teaching-learning situation. This does not augur well for the eventual success of the reform, irrespective of how good and well intentioned it may be.

9.1 The Change Process

9.1.1 Recognising that Change is a Complex Process

The traditional classroom realities depicted in this study, which contrast sharply with the intentions of local policy makers, bear witness to the fact that achieving change in teachers is neither linear nor smooth. To start with, change is a process not an event (see Fullan, 2001). More precisely, it “is a complex process which involves the head and the heart, the personal and the professional” (Day, 2000, p. 126). As such, it does not happen there and then when ‘someone from above’ demands it. The fact that well-established teacher practices are hard to change (Calderhead, 1987) probably renders the whole process even more difficult. Change is in reality a ‘risky journey’ that is subject to failure (Black et al., 2003). Bringing changes in assessment does not appear to be an exception. In fact, whilst a growing body of research documents the difficulty of introducing effective formative assessment into classroom practice (William & Black, 1996), the introduction of ‘Teacher Assessment’ in England and Wales has even led, amongst others, to the ‘evidence gatherers’ model that largely neglects the

formative aspirations of the reform (see Gipps et al., 1995). Such difficulties are to be expected once one realises that, for instance:

Implementing assessment for learning requires personal change. It means changing the way a teacher thinks about their teaching and their view of their role as a teacher. Since the way a teacher teaches is inextricably linked with their own personality and identity, ultimately it means changing yourself. (Black et al., 2003, p. 80)

I can moreover add from personal experience that, apart from overcoming this ‘Zeitgeist of a previous age’ (see Broadfoot, 1995), the successful implementation of assessment for learning requires one’s personal ‘inner change’ to translate itself into an outer, visible change. For only when change becomes part of the natural behaviour of the teacher can it be said to be successful (Hopkins, 1994). My less-than-successful efforts as a teacher to move across the continuum from inner to outer change suggest that it is not enough to have what Hargreaves (1994) defines as the ‘capacity and desire for change’. Instead, with assessment clearly being a context-bound activity (see section 1.2.2), the operating context must necessarily come to the fore of the assessment change debate. In reality, every teacher’s deep, influential sense of practicality – which is the distillation of purpose, person, politics and workplace constraints – leads to decisions based on what works in his or her particular context (Hargreaves, 1994). To implement change in teachers is therefore to ask much more than whether what is being proposed works. It is also to ask whether it fits the context, whether it suits the teachers, whether it is in tune with their purposes, and whether it helps or harms their interests. The sheer number of factors at play indicate the complexity involved when pursuing change – a case against approaching change with the mentality that there are hard-and-fast rules that apply universally (Fullan, 2001).

9.1.2 Recognising that Change is Possible

The understanding that what works in one situation may or may not work in another explains why “it is not possible to have all the answers in a ready-made package” (Weeden et al., 2002, p. 38). But whilst general guidelines to implementing change do exist, research findings on the change process should be used less as an instrument of ‘application’ and more as a means of helping those involved to ‘make sense’ of their own situation so that they may be in a better position to move things forward (Fullan,

2001). The idea is to enable teachers and others engaged in education to improve their ability to see and think about what they do (Eisner, 1985; cited in Woods, 1990). This has indeed been the guiding principle of Black et al.'s (2003) assessment project that set out to explore changes in teachers' classroom practices that can make teaching and learning more effective. As a matter of fact, the teachers involved participated throughout the project in in-service education and training events that gradually came to reflect their agenda, and benefited from classroom visits by members of the research team that helped them to reflect on their assessment practices. These teachers subsequently reported that although the changes in them happened slowly – almost imperceptibly – when they looked back, the cumulative effect had been substantial.

Albeit Black et al.'s (2003) project took place in a different context from that of the present study, it still embodies the exciting possibility that things can change, even if slowly, when there are concerted, genuine efforts. But even here, the number of teachers who, for some reason or other, had to abandon their project amply demonstrates that success can never be guaranteed. Again, although Black et al. (2003) set out to study how different teachers might realise, within the normal constraints of the curriculum and external testing requirements, substantial improvements in the learning of students through the development of formative assessment, their project report is not as 'comprehensive' as this may suggest. It does not tackle, in effect, a number of important issues that are pertinent to my study. To mention but a few, their report throws no light on what would happen when teachers fail to find support within and outside schools, when teachers are not already somewhat aware of assessment problems and are thereby prepared and committed to take on fresh initiatives, and when teachers are actually preparing students for high stakes external assessments. Notwithstanding this, I still consider this project as an important manifestation that research (in this case, the general principles of assessment for learning) can be put into practice when circumstances permit. It is now my turn, both as a researcher and as a teacher, to explore ways of achieving improvement in TCAP (and ultimately in students' learning) within the present context that is replete with situation specific difficulties and constraints. My understanding that this would entail complex changes is comforted by the knowledge that, as Fullan (2001) points out, contrary to simple changes that are easier to carry out but may not make much of a difference, complex changes demand more effort but promise to accomplish more.

9.2 Looking beyond Teachers' Explanations

9.2.1 Revisiting Teachers' Explanations

My colleagues tend to explain their classroom assessment practices as the end result of what they believe in and know about assessment under the influence of what their circumstances permit (see section 7.7.3). Whilst their beliefs and knowledge of assessment practically mirror their views on teaching and learning as being respectively 'transmission' and 'practice' (see section 7.5.2), the assessment constraints they identified can be classified almost exclusively under the school and national contexts (see section 7.7.2). It transpired further that the assessment constraints reported by these teachers seem to affect primarily the 'quantity' rather than the 'quality' of their assessment practices (see section 7.7.4). So much so that when asked to comment on how their assessment practices would change in an ideal constraint-free environment, they appeared to be mostly interested in consolidating their existing practices. In fact, even if some of them spoke about the possibility of introducing some assessment 'novelty' or other on the side (see Table 7.12), they all practically anticipated to do 'more' (which was invariably presented as 'better') of what they are already doing. These ideal scenario 'projections' by my colleagues – which suggest strongly that their practices would not change much in a constraint-free environment – are in line with their rather positive outlook on the quality of their assessment practices (see section 7.7.1). These teachers' apparent lack of desire for a drastic change in assessment should circumstances permit signals in turn a high level of conformity between their assessment beliefs (i.e., the manner in which they consider that things ought to be done) and practices (i.e., what they actually do).

9.2.2 A Critical Look at Teachers' Explanations

Speaking of my own assessment practices, I would say that these are moulded by my assessment beliefs and knowledge, and the constraints in which I operate. It is my recognition that these constraints also rest 'within me' that mostly sets me apart in this respect from the other participants. Indeed, I readily recognise that the 'external' constraints identified by my colleagues are equally restricting on my practices (even if I would certainly add to their list – see section 9.3). On the other hand, the manner in

which they view classroom assessment – characterised as it is by an emphasis on measurement, management and accountability – reminds me a good deal of myself before I set out to embrace assessment for learning as widely and as effectively as I can. The fact that I still basically use the same assessment practices as the other participants amply indicates how delimiting the constraints are. It is as if my colleagues and I manage to arrive at practically the same ‘spot’ albeit following different ‘tracks’. To a certain extent, the main difference between us may be that whilst they basically assess in harmony with their beliefs – which explains why their ‘state of readiness to engage in additive or transformative change’ (see Day, 1999) is low at best – my beliefs and actions are largely in conflict. This conflict lies behind my desire for change – which is reflected in my willingness to explore avenues as to how my practices can be brought in line with my beliefs.

Bearing in mind that change can be made most effectively when it is synchronised with teachers’ own desires for it (Hargreaves, 1994), my colleagues’ almost non-readiness for assessment change introduces an important dimension into what actually impedes the improvement of TCAP. For it appears that, in addition to the mostly ‘external’ explanations that they put forward, equally delimiting constraints may actually be at work within the teacher him or herself. The point is that teachers, being fundamental intermediaries to the change process, can either spur on or halt assessment reform initiatives according to the manner in which they ‘judge’ the situation in their charge. Day (1999) presents a strong argument for this:

Teachers are willing to engage in change according to whether they perceive a need, diagnose a problem, and conceive of a response to the problem that is both within their intellectual and emotional capacity, and appropriate to their personal, educative and ideological perspectives and the context in which they work, and have access to support. (p. 100)

It thus appears that teachers’ engagement in assessment change is subject to two overriding dimensions – namely, the inner or personal dimension and the outer or external dimension. To nurture one and not the other, as in my case with the inner dimension, possibly only serves to create false hopes for change resulting in a deep sense of frustration. Likewise, to name one as a source of constraint and not the other, like my colleagues have largely done with the outer dimension, is to deny the complexity involved in the assessment process. As teachers are social learners, their

actions are consequently moulded by both their individual capacities and a deeply influential sense of what works within one's own particular context (see Hargreaves, 1994). This understanding makes it unsustainable to argue, as my colleagues seem to imply, that TCAP can be improved by simply creating the 'right' contextual conditions (i.e., unless one is to include oneself as part of the context). The manner of little consequence in which they proposed to modify their assessment practices in a constraint-free environment demonstrates how improvements in the operating context without parallel initiatives aimed at helping teachers to change internally hardly offers any guarantees. Again, had I not 'grappled' with issues in educational assessment, I would most probably neither have become aware of my assessment limitations nor would I have sought out remedies – and this irrespective of what was 'out there'.

9.2.3 Identifying Three Levels of Concern

Our collective experiences thus show that improvement in TCAP call for action within two distinct, but highly interrelated, dimensions – what Day (2000) calls the 'working context' and the 'personal context' of the teacher. Action in the personal dimension is needed so that teachers may become more sensitised to the principles and practices that underline the new assessment paradigm, and action in the working dimension is needed so that the contexts in which teachers operate would become a vehicle for change rather than remain an obstacle to it. These personal and working dimensions, in reality, span over three levels – namely, the 'teacher' level that corresponds to the personal dimension, and the 'school' and 'national'¹² levels that correspond to the working dimension. These three levels are described below:

- The teacher level: This refers to the teachers themselves – their knowledge base and professional development, their beliefs and practices, and their work ethics.
- The school level: This refers to people, structures, policies and practices inside schools (e.g., administrators, colleagues, students, regulations and practices) that are in a position to influence and shape the manner in which teachers operate.

¹² Given Malta's small geographical size and highly centralised educational system (see Zammit Mangion, 1992), one is practically at national level just beyond one's own school. Unlike much bigger countries, we do not generally speak of local or regional levels as far as education is concerned.

- The national level: This refers to people, structures, policies and practices outside schools (e.g., teacher educators, politicians, policy makers, examination boards, parents, and employers) that are in a position to influence and shape what goes on inside them and thus ultimately what teachers do.

Like the Russian nesting dolls, each of these three levels is tightly compressed within and/or around the others. Given that these levels work interactively (see Fullan, 2001), it follows that improvements in TCAP would require each level to be carefully analysed in order to build on the identified strengths and to try to eliminate the identified weaknesses. The ultimate aim is to construct an all-encompassing ambience that helps assessment truly become an integral element of the learning process.

9.3 Acknowledging the Difficulties

This section highlights the difficulties – which spread across the ‘teacher’, ‘school’ and ‘national’ levels – that hinder the implementation of the underlying principles of the new assessment paradigm inside PMI classrooms at the University of Malta’s sixth form college. These are consequently context specific difficulties to assessment change that cannot be easily generalised to other teachers, other schools and other countries. The same would apply to finding ways of redressing these difficulties. For I find that Black et al.’s (2003) remark that different teachers start at different points and follow different change trajectories can be extended to include different schools, regions and countries. The ensuing identification of these difficulties serves primarily to delineate the challenges ahead – that is, to establish what my colleagues and I, our school, and the Maltese educational system need to address in order to move towards a classroom assessment scenario that is more conducive to learning than at present.

9.3.1 At Teacher Level

My colleagues’ overall views on teacher education come closest to Elliott’s (1993b) ‘social-market view of teacher-education as a production/consumption system’. Within this essentially behaviourist perspective, the initial training phase of teachers consists of a short induction – just enough to ensure that the new teachers acquire a few basic behavioural skills that are sufficient for them to function within the school.

The emphasis that my colleagues lay on experience in teaching reflects in turn how the continuing education phase of teachers from a social-market perspective is concerned with progressively developing higher-level skills over a longer period of time. Professional learning from this perspective is an outcome of training rather than education. My colleagues tend to equate such training to ‘learning on-the-job’ – which explains their little faith in preservice and inservice professional programmes, and also their general lack of interest in developing an interactive relationship between theory and practice. As a result, published educational research appears to have little or no influence on their classroom practices. In addition to this ‘gulf between research and practice’ – which is a common phenomenon amongst teachers in general (Mitchell, 1999) – my colleagues emerge as primarily concerned with retaining their classroom autonomy and with no real interest in collaborating with one another. Although we all largely operate within what Gipps et al. (1995) call ‘autonomous isolation’, the other participants also tend to accept their own classroom practices uncritically and to keep their classroom practices, once established, for good.

With these characteristics, my colleagues clearly do not conform to Hoyle’s (1980; cited in Woods, 1990) description of ‘extended professionals’. For extended professionals – apart from being theory-friendly, concerned about educational and teacher development, and see teaching as a rational activity amenable to improvement on the basis of research and development – are also concerned with locating their classroom teaching into a broader educational context, comparing their work with that of other teachers, evaluating their work systematically, and collaborating with other teachers. Neither does the teacher reality depicted in this study mirror the notion of ‘reflective practitioner’ developed by Schön (1983) that, as Day (1999) holds, has become synonymous with ‘good’ practice. Schön’s ‘reflective practitioner’ is essentially innovatory and creative, discovering problems and issues, inventing and experimenting in the search for solutions, and continuously adopting as he or she reflects ‘in’ and ‘on’ action. This stance – which effectively celebrates teacher artistry (see Harris, 1998) – benefits teachers’ problem solving efforts. Indeed, with many of their ‘difficulties’ being “uncertain practical problems that require unique and idiosyncratic approaches to solution because of their strong ties to specific contextual factors ... the skill required is that of intelligence and artful orchestration of knowledge and technique” (Clark & Yinger, 1987, p. 98; cited in Woods, 1990).

From these observations, it appears too simplistic to argue that my colleagues' lack of assessment education, albeit an undeniable reality that contributes to undermine the link between assessment and learning in their classrooms (see section 8.3.1), is the main difficulty at teacher level. It would be more sensible to consider their lack of assessment literacy as part of a 'bigger' problematic reality. Indeed, the manner in which these teachers live classroom assessment has essentially emerged as a manifestation of 'what they are and stand for' – for example, viewing teaching as transmission and learning as practice, viewing the teacher as in control, valuing personal autonomy and shunning collaboration, seeing teaching as a craft that requires little or no critical reflection, linking personal success to the examination success of their better students, lacking the motivation to pursue professional interests, and so on. The challenge is thus to address this 'bigger' reality – which basically relates to teachers' knowledge base, beliefs, values and attitudes – rather than the technical issues of assessment in isolation. It is simply not enough to introduce teachers to new assessment techniques and subsequently 'convincing' them about their use. For assessment changes 'forced' on teachers against their beliefs will be superficial and probably short-lived (see Day, 2000; also Black et al., 2003). Real change requires instead teacher beliefs to be significantly altered (Fullan, 2001).

9.3.2 At School Level

My colleagues and I concur that a number of school factors – which include large student groupings, lack of classroom space, limited contact time with students, and the end-of-first-year examination (see Table 7.11) – constrain our classroom assessment practices. But then, although my colleagues also referred to a number of other factors that 'bother' them at school level – such as, not being involved in the running of the department, the presence of weak and unmotivated students, the lenient promotion system, and the lack of grading guidelines – they do not perceive them as constraints on classroom assessment, at least not with regards to their own practices. However, as the following two examples reveal, there is evidence to suggest that such factors can in fact influence – with obvious negative consequences for learning – the manner in which these teachers relate to assessment. First, many colleagues seem to react to the school's policy of enrolling minimally qualified students by largely ignoring the 'weaker' ones throughout the various phases of the lesson, assessment included.

Second, a couple of teachers claimed that one of the reasons for giving demanding class tests (which subsequently lead to lower Assessment Marks – see section 7.3.3) is to compensate for what they judge as a lenient school promotion system. In both cases, these practices work against the emphasis on learning embedded in the new paradigm. For whilst in the first instance assessment is not intended to benefit all students, in the second instance assessment focuses on selection rather than learning.

These two examples highlight moreover how teachers can strategically redefine their work so that it remains underpinned by their own values and identity (see Day, 2000) – in this case, in order to compensate for what my colleagues see as ‘deficiencies’ in our school system. This ‘redefinition of work’ basically entails the sidelining of those students whom they see as inconsistent with our school’s pre-university orientation (see section 8.2.1). It is ultimately a means for these teachers to cope and survive within the system, their way of gaining some control over how they carry out their job. The realisation that teachers can ride in this manner over ‘problems’ thrown at them by schools – what Woods (1990) calls ‘accommodation’ – gives a clear signal that teacher practices are also influenced by how teachers perceive their school context. This means that, over and above any teacher-identified constraints, the school context itself needs also to be addressed from the wider perspective of its impact on teachers. This recognises that the school context is as constraining as the number of ‘bad’ practices that it induces, directly or indirectly, in teachers.

The understanding that what teachers do (or fail to do) is somewhat conditioned by their school environment implies that schools with ‘non- or mis-educative environments’ (see Cole, 1997) are particularly problematic. One such problem in this study is the inexistence of a school assessment policy (as different from the existing promotion criteria – see section 2.4.4). It is not enough that schools leave it to individual teachers to juggle on their own the conflicting demands of bureaucracy and learning – they need instead to provide a coherent set of principles and procedural knowledge about assessment (Biggs, 1996). Another problem is the lack of what Nias et al. (1989) call the ‘culture of collaboration’. My colleagues and I, instead of working as a team within an environment that has few status differentials and where we accept and foster our interdependence while valuing individuality in an atmosphere of security, live in a state of ‘balkanisation’ (see Hargreaves, 1994). A balkanised

work environment is characterised by teachers in ‘cubbyholes’ (see section 6.3.4) – an ambience that severely restricts collegiality and collaboration amongst them, creating instead competitive territoriality and lack of opportunities for teachers to learn from and support one another. The fundamental problem here – and which also emerged in this study (see section 9.3.1) – is that such school conditions generate feelings and psychological states that militate against reflective practice and professional growth (Cole, 1997). This clear link between the difficulties at teacher and school levels signals that the responsibility for the quality of education (of which assessment is an integral part) is also a matter for the school, not just the individual teacher (see Grundy, 1994; cited in Day, 2000). For, in reality, it is only after the underlying difficulties at school level are addressed that the teachers will be able to engage freely and meaningfully in the kind of reflective practice and professional development that brings meaning to their own lives and the lives of their students (Cole, 1997).

9.3.3 At National Level

With schools mediating societal demands (Woods, 1990), it is quite clear that the action at teacher and school levels advocated in the previous two sections needs support from outside schools. My colleagues appear to link such external support simply with ‘improvements’ in the MATSEC certification system. For, at national level, they feel constrained in their assessment practices by the lowering of standards at O-Level, the too demanding matriculation system, the relatively vast PMI syllabus and the examination questions (see Table 7.11). A deeper reading of the situation however suggests that the certification system – under its various facets, not just those mentioned by my colleagues – is but one of the issues that need to be addressed at national level. As with the teacher and school levels, it is not just what the teachers studied had to say about the national level that needs to be taken into consideration, but also the manner in which this level impacts on their practices. This exercise would eventually indicate the needed changes outside schools so that the newly aligned personal, school and national contexts would bring TCAP closer to learning.

For a start, the very nature of the local educational system – characterised as it is by a rigid, monocultural, centrally determined curriculum that is policed by a narrow examination regime – is unlikely to provide the circumstances in which innovative

approaches to teaching, learning and assessment will flourish (see Torrance, 2000). Further to this, Hargreaves' (1984) three broad areas of constraints over teachers that originate outside schools also appear to exert a negative impact on the participants' practices. The first constraint concerns what Hargreaves sees as a 'fundamental contradiction' between the two major goals of the education system. For schools and teachers are expected to educate all students to their maximum potential and to give individuals their due, but at the same time they have to select and socialise students, primarily through examinations, for society. According to Hargreaves, the difficulty of reconciling these two goals tempts schools and teachers to channel their whole organisation and pedagogy towards practices that favour examination success at the expense of their other more educational responsibilities. True to this, the trend amongst the teachers studied is to focus on examinations – be it whilst teaching (see section 6.4) or assessing (see section 8.1.3) – within a school scenario in which the administrators' interest in students' learning appears to be primarily linked to knowledge of Assessment Marks and examination results.

The second constraint that Hargreaves (1984) mentions is the level of funding. He argues that material constraints curtail the school reality in a number of ways – for instance, it determines the number of teachers employed, the length of students' schooling, the kind and state of school buildings, the provision of books and other teaching aids, the amount of in-service teaching, and so on. The funding difficulties of the school in this study (see section 2.4.2) are indeed reflected in the high student-teacher ratio, lack of teaching resources and professional development opportunities, inadequate classroom space and amenities, overcrowded teacher offices that lack proper facilities, and much school energy being spent on money generating activities as opposed to educational ones.

The appearance from time to time of certain educational ideologies, the so-called 'good practices' upon which teacher career advancement depends, constitutes the third of Hargreaves' (1984) constraints. This is constraining in the sense that teachers have the problem of meeting these practices along with the other goals of the system, some of which may be in tension with the new ideologies or practices being proposed. This study shows that the manner in which teachers react to this externally imposed situation can have important repercussions on classroom assessment. For instance,

when the teachers studied choose to ignore the ‘directive from above’ to give lectures (see section 2.4.3) and insist on teaching (see section 6.5.1), their relatively ‘closeness’ to students benefits classroom assessment. On the other hand, my colleagues’ tendency to add on to their more than sufficient mathematics content qualifications rather than invest in their professional development as teachers (see sections 6.2 and 6.3.5), albeit progressively beneficial to their pay cheque (see section 2.4.1), is not likely to improve their classroom practices, assessment included.

9.4 Exploring Prospects for Improvement

It is now time to explore avenues that would facilitate the implementation of the new assessment paradigm inside our classrooms in the knowledge that this is hindered by difficulties at teacher, school and national levels. In search of this, one has however

to contend with both the ‘what’ of change and the ‘how’ of change. Meaning must be accomplished in relation to both these aspects. It is possible to be crystal clear about what one wants and be totally inept at achieving it. Or to be skilled at managing change but empty-handed about which changes are most needed. (Fullan, 2001, p. 8)

The understanding that the ‘what’ and the ‘how’ constantly interact and reshape each other (Fullan, 2001) indicates the high level of contextualisation involved, both with regards to space and time. This acts as a clear warning against drawing unfounded generalisations from the conclusions and suggestions made here.

9.4.1 The What of Change

My vision for classroom assessment is based on a view of the learning of mathematics as a socially constructed process, not a fixed hierarchy of skills and concepts to be mastered. It follows that the assessment system – in recognition that mathematics is a dynamic set of interconnected, humanly constructed ideas – must allow students to engage in rich activities that include problem solving, reasoning, communications, and making connections (Romberg & Wilson, 1995). During these activities,

Teachers have to refrain from an imposing role, and to take on a role that includes helping students develop productive small-group collaborative relationships,

facilitating mathematical dialogue between students, and, above all, orchestrating a discussion around issues that are significant in view of the envisioned learning trajectories. (Seegers & Gravemeijer, 1997, p. 263)

I see in this pedagogical approach – which moves away from that in traditional classrooms in which “it’s the teacher’s to know and the student’s to find out” (Seegers & Gravemeijer, 1997, p. 262) – an opportunity to widen the base of mathematical activity inside the classroom to cover the three levels of assessment mentioned by de Lange (1995) (i.e., lower, middle and higher – see section 3.1.2). At the same time, the understanding that learning is like “an image that is gradually brought into sharper focus as the learner makes connections ... with specific bits of knowledge situated within some larger design that is continually being reorganised or redesigned in an organic manner” (Romberg & Wilson, 1995, p. 5) precludes the use of an atomistic model for assessment (see Gipps, 1994). Indeed, it was partly in response to assess the level of understanding and the complexity of understanding that the alternative assessment paradigm came into being. Classroom assessment must however also reflect the manner in which the world has moved on. Nowadays “basic skills are not enough – the majority of the population, not just the elite, needs to become flexible thinkers, reasoners and intelligent novices” (Gipps, 1994, p. 161). This parallels Romberg and Kaput’s (1999) integrative view of mathematics that stresses the acquisition of understanding by all to the highest extent of their ability rather than the selection and promotion of the elite. A prerequisite for this is the development of an assessment system that fosters growth towards high expectations and supports high levels of student learning (see NCTM, 2000).

9.4.2 The How of Change

I now move on to suggest how the classroom assessment reality outlined above can be achieved through action at the three levels of concern identified previously. This action is based on the understanding that ‘solutions’ to complex problems (such as the one at hand) require what Senge et al. (2000; cited in Fullan, 2001) call a ‘learning orientation’. This means involving everyone in the system (irrespective of whether they are inside or outside schools) in expressing their aspiration, building their awareness, and developing their capabilities together. This orientation develops the ‘shared meaning’ that is so essential to successful change (see Fullan, 2001).

9.4.2.1 At Teacher Level

The focus for action here is to capacitate teachers. This would serve teachers to better appreciate the intimate links between teaching, learning and assessment, and also to understand that an integrated approach between these three pillars of classroom life requires a new conception and forms of assessment. But this transition from the traditional to the new assessment culture cannot be seen as an add-on process. The route to change is instead through first creating a desire for it (see Hargreaves, 1994). As effective classroom assessment depends on an understanding by teachers that learning is a sense-making process that students need to experience for themselves (Bright & Joyner, 1998), assessment education programmes for teachers have to be conceived so as to mirror and support the wider picture of their professional education. In particular, all efforts aimed at making beginning or practising teachers assessment literate must necessarily be built around current conceptions of constructivist pedagogy. This would make it more likely for TCAP to be ‘integral’ to their ideology and to their other classroom practices – an eventuality that increases the chances for classroom assessment to have a positive impact (see Broadfoot, 1995).

If teacher assessment education is to leave its mark, it needs to stimulate teachers’ professional growth in additive (by taking knowledge, skills, and understanding a step forward) and transformative ways (by leading to major changes in beliefs, knowledge, skills and understandings) (see Day, 1999). Professional development, apart from being sustained, needs also to be timed to ‘fit’ the needs of teachers in relation to their phase of experience, career development and lifelong learning cycle, as well as the needs and demands of the system in which they operate. In particular, given the teachers’ powerful sense of practicality (see Hargreaves, 1994), it would be unyielding for such programmes to prescribe a model of effective classroom assessment action. It is far better for these initiatives to deal only with the general principles of ‘good’ assessment, and then to let teachers find their own ways of implementation (see Black et al., 2003). The understanding of teacher professional action as the power to make independent judgements and to exercise personal discretion, initiative and creativity through work (Schön, 1983) suggests that pre-service and in-service courses should serve primarily to provide stimulus, challenge and support for the work of teachers in inventing new practical knowledge about classroom work, assessment included.

When teaching is viewed, as Elliott (1993a) suggests it is best, as a practical science in which the relationship between theory and practice is interactive, Stenhouse's (1985; cited in Woods, 1990) plea for teachers to be educated to develop their 'art' aligns teacher education to the 'teacher-researcher' and the 'action research' movement. The emphasis of this perspective on 'grounded data' (see Day, 2000) challenges the traditional treatment of theory and practice as separate entities in teacher education.

Theory is implicit in practice, and the relationship between theory and practice in teacher education is not one of implementation – theory being translated into practice – but a continuously interactive one. ... Theory can provide the analytical and conceptual apparatus for thinking about practice, while practice can provide the opportunity for the testing and assimilation of theory. (Calderhead, 1988, p. 9)

This interrelationship between theory and practice in terms of knowledge as embedded in rather than existing outside action is at the heart of Schön's (1983) notion of reflective practitioner. Seeing theory as derived from practice that informs its further development legitimises teaching as a knowledge-based, intellectual activity in which teachers are not only capable of deconstructing but also of reconstructing experiences. In recognition that current conceptions of teacher professionalism are no longer grounded in notions of esoteric and specialist expertise, initiatives aimed at capacitating teachers need to focus on enabling teachers to develop their critical and creative powers. This calls for helping teachers to cultivate their professional judgement that then acts as a highly significant filter in interpreting the significance of research evidence for particular students and classroom contexts (Pollard, 2002).

9.4.2.2 At School Level

The focus for action here is to turn the school into the 'centre' of change (see Hopkins, 1994). It follows that the school needs to function as a learning organisation in which teachers are the leading learners (see Day, 1999; also Southworth, 1994). This would create an in-house supportive environment, in which teacher creativity is valued and encouraged, that enables teachers to maintain and build upon an assessment purpose and vision nurtured through their professional education and a lifelong commitment to learning. Such an environment however requires a participatory-democratic school ethos as opposed to an authoritarian one (Woods, 1990). The point is that

improvement in the quality of education cannot materialise without also changing the school organisation (Hopkins, 1994). Educational leaders must consequently learn to lead, not from the top of the traditional pyramid of authority in schools, but from the centre of a web of interpersonal relationships, with people rather than through them (Bezzina, 1997). This would help to establish the 'learning partnership' between teachers, students and parents desired by Day (1999). Within this reality, the classroom becomes more of a community in which the teacher and students have different roles that need not all be age or experience related (Gates, 2002). It would moreover be normal in this partnership for the more skilled, knowledgeable and experienced teachers to mentor their novice colleagues, for teachers to be involved in school examinations, for teachers and students to 'dialogue' outside the lesson, for teachers to see colleagues at work and recognise their efforts, for well-informed parents to use their enormous potential for facilitating their children's learning (see Filer & Pollard, 2000), and ultimately for celebrations of success.

From this perspective, the school cannot simply be construed as the location where teachers work. It has instead to play its part in their education. In particular, as the responsibility for the quality of education is also a matter for the school, not just the individual teacher (Grundy, 1994; cited in Day, 2000), the school has a role in classroom assessment that goes far beyond providing the right personnel and the necessary logistics and material resources. Other than embracing an enlightened assessment vision and policy that rightly balances its administrative needs with the learning needs of its students, the school has to work towards rooting within it a collaborative community of teachers (see Nias et al., 1989). This culture – which does not preclude the organisation, whenever the need arises, of school-led professional development assessment courses (see Day, 1999) – will however only flourish once any existing 'balkanisation' (see Hargreaves, 1994) within the school is replaced by collegial discussion within 'communities of practice' (see Elwood & Klenowski, 2002). Participation in such communities "is essential for clarifying understandings of the purposes and the practices involved, for ensuring a context of collegial support as individual teachers take on the risks of changing their classroom practice, for promoting reflection on the issues raised and in providing the forum for evaluation" (Black et al., 2003, p. 115). Should these communities develop into 'communities of inquiry' in which teachers cultivate inquiry approaches to their practice and together

use inquiry approaches to develop their practice, teachers would then have the opportunity to experience a reflexive relationship between inquiry and development that helps them to grapple better with the complexities of teaching (Jaworski, 2004).

In recognition that such communities strengthen the capacity for change, school administrators would do well to create ongoing opportunities for teachers to come together in this manner (Black et al., 2003; Bright & Joyner, 1998; Steadman, 1998). As a matter of fact, this teacher-to-teacher ‘dialogue’ serves to build an expertise in assessment that is truly unique to the teachers involved (Black, 1998). But just as it is now recognised that professional autonomy needs to be reconceptualised in terms of the professional community of the school (Grundy, 1994; cited in Day, 2000), it is also vital that the school facilitates collaboration and collegiality amongst teachers without controlling it (Hargreaves, 1994). For requirements for teamwork and collaboration that appear to be eliminating opportunities for independence and initiative (which constitute the heart of professional action – see Schön, 1983) may actually result in unhappiness and dissatisfaction (Hargreaves, 1994).

9.4.2.3 At National Level

The focus for action here is to develop an outside school context that is not only amenable to change, but also acts as a catalyst for change. For even though the school may become a ‘centre of change’, it is still “embedded in an educational system that has to work collaboratively and symbiotically if the highest degrees of quality are to be achieved” (Hopkins, 1994, p. 78). At its most basic level, any commitment to change and improvement that comes from outside the school is bound to involve financial considerations. As a matter of fact, the more money that is put into teaching resources, the better the buildings and equipment, the greater is the range of opportunities (Woods, 1990). Society needs however to also invest in teachers if it wants to develop schools (Day, 1999). Such external support would help teachers to experience school as development rather than survival (see Cole, 1997), and to see teaching and all that comes with it as challenges rather than burdens to be borne.

Particularly at issue here is the kind and quality of the training and development opportunities – in which assessment needs to gain centrality (Black, 1998) – that

teacher education institutions offer teachers throughout their career. In order to enculture students into mathematics, these opportunities must help teachers understand that “It is not enough to teach them mathematics, we need also to educate them *about* mathematics, to educate them *through* mathematics, and to educate them *with* mathematics” (Bishop, 1991, p. 3; emphasis in original). This mathematical enculturation requires in turn an assessment education that promotes a system fundamentally concerned with ‘a way of knowing’ that is a socially constructed set of ideas and meanings, which is reconstructed in turn by each new generation (see Bishop, 1991). This kind of professional preparation would make it less likely for teachers to emulate a pedagogy that focuses on examination success at the expense of not only distorting the nature of mathematics, but also of impeding the students’ acquisition and use of other mathematical knowledge (see Schoenfeld, 1988).

On their part, policy-makers would do well to heed to Hopkins’ (1994) advice to avoid a ‘top-down’ approach to change, as efforts aimed at implementing change through mandated practices are notoriously known for failure (see Day, 1999; Eisner, 1993; Hargreaves, 1994). Probably even less effective than this would be to impose policies through fears of public ‘naming and shaming’. The crucial point is that external reforms need to be sensitive to the situation in individual schools. Hopkins (1994) recommends the use of ‘integrative implementation strategies’ to meet this challenge:

This implies a linkage between ‘top-down’ and ‘bottom-up’ [approaches] ... Ideally, ‘top-down’ provides policy aims, an overall strategy and operational plans; this is complemented by a ‘bottom-up’ response involving diagnosis, priority goal setting and implementation. The former provides the framework, resources and a menu of alternatives; the latter provides energy and school-based implementation. (p. 79)

But it is then equally important that individual teachers in particular schools are given the space to work out the details of this implementation in their own classroom (see Black et al., 2003). Implementation at classroom level is however still at risk when a tightly specified national curriculum, backed by high-stakes assessment procedures, diminishes the opportunity for teachers to respond to student needs. Such a reality is not only unlikely to promote meaningful learning, but also produces exclusion with disengaged and disaffected students simply withdrawing in their minds and/or bodies (Pollard, 2002). Given the significant influence that the certification process exerts on

classrooms (see section 3.0), examination bodies can support reform in classroom assessment by working towards making their assessments fairer for all concerned (see Chetcuti, 1998; Gipps & Murphy, 1994). In particular, as there are no fundamental reasons as to why teacher assessment cannot play a role in external assessment (Black, 1998; Calfee & Masuda, 1997), its inclusion in certification procedures – apart from contributing to make this process fairer (see Broadfoot, 1996; Broadfoot & Black, 2004; also Desforges, 1989; NCTM, 2000) – can actually ‘push’ teachers to become interested in and to learn more about the various facets of assessment.

9.5 Concluding Comments

The research journey outlined in this study has basically explored the classroom assessment practices of PMI teachers in my own school by seeking and providing answers to the four research questions identified in section 1.3.1. My understanding that these TCAP are not particularly supportive of learning and my consequent call for change build on Weeden et al.’s (2002) vision embodied in the prospect of all students being enthusiastic and effective learners. This vision contrasts with the current emphasis on pushing some of the ablest students to reach the highest possible standards whilst leaving the vast majority of students to achieve well below their potential (see Boaler et al., 1998). Change is needed as it makes no sense for the largely ‘not for learning’ scenario identified in this study, which is particularly unfair on the so-called ‘less academically gifted’ students, to continue flourishing. Even more so in the knowledge that the overall picture presented here suggests that other local schools dominated by traditional teacher beliefs and practices, top-down management, lack of vertical and horizontal dialogue, and inexistent opportunities for teacher professional development are equally likely to produce practices that better serve the managerial than the professional functions of assessment.

Whilst I am under no illusion that change can be achieved overnight, my strategy at teacher, school and national levels for improving classroom assessment (see section 9.4.2) recognises further that the formulation and legislation of policies pushing for assessment reform is not sufficient (see section 9.0). On the contrary, even though progress is very hard to achieve without the needed adjustments at school and national levels, there is no denying that all is likely to be for nothing should one fail to

capacitate teachers in the hope that their aroused desire for change is followed by action (see Hargreaves, 1994). From whichever angle I look at the situation investigated in this study, the first priority always appears to be that learning gains or regains centrality in teachers' discourse and action. Without wanting to diminish the action that needs to be carried out subsequently with teachers and at school and national levels, I firmly believe that this is where the key to improved classroom assessment actually lies. For once the teachers become committed to teaching for learning, they are more than likely to accept the philosophy and accompanying practices of the new assessment paradigm, and to actively lobby for an environment at school and national levels that permits and facilitates its realisation. The transformation of teachers from mere system survivors to a catalytic force for change, which would certainly benefit the reprofessionalisation of teachers, is not however an easy task. It is instead a learning process that can create feelings of doubt, uncertainty and, sometimes, a loss of confidence in some individuals. To counterbalance these emotional costs – which constitute the painful side of learning – the school needs to become a person-centred, sensitive and caring institution (Southworth, 1994).

The complexity involved in moving from 'where we stand' to 'where we should be' highlighted in this study suggests that the relatively unproblematic implementation of 'assessment for learning' inside classrooms described by Black et al. (2003) does not readily apply when the teachers and the schools involved are not as receptive to change as the ones in their study. Moreover, as Torrance (2000) argues, it may be that issues of complexity and difference in assessment (he refers in particular to the unsatisfactory development of Records of Achievement in the UK) will eventually lead to a radical rethinking about educational assessment. It would however be a mistake to simply focus on certain aspects of the reform that have been tried and found wanting in order to justify the turning back of the wheel. Whilst prudence always pays because "change and improvement are not necessarily synonymous" (Hopkins, 1994, p. 74), we also need to appreciate that the rewards for persevering in this direction appear to be worth waiting and fighting for. And when the general public begins to understand and value these alternative forms of assessment, not only would this help to improve schooling for students, but it would also contribute to a broader, more generous conception of education itself (Eisner, 1993).

My Journey as a Teacher Continues ...

My research journey came to an end at school after I re-established myself as a 'teacher only' within the mathematics department (see section 5.3.4). Although some years have now already passed, my teaching future remains linked, at least for the foreseeable future, to the same school where I carried out the research. This is where the educative aspect of my research comes in (see Hammersley, 2003). As 'Michael the teacher' I have always felt the urge to utilise my gained expertise in educational assessment to help create an environment that is conducive to improvement in classroom assessment practices in order to raise the standard for all (see Weeden et al., 2002). My efforts in this direction have targeted three different, but highly interrelated, levels – the personal (to which I have already referred), the school, and the national levels. In particular, I have worked on the school and national dimensions in my capacity as a teacher who is genuinely interested in collaborating with all colleagues, not just those who participated in this study, towards affecting change. Believing that it is better to think big, but start small (Fullan, 1988; cited in Hargreaves, 1994), I consciously followed two parallel, but complementary, strategies. First, to work towards the creation of an assessment dialogue amongst colleagues that encourages us to reflect upon and to challenge our current practices – what Farrugia and Attard (1989) call 'the reorientation of teachers'. Second, to get involved in departmental- and school-based initiatives that have the potential to render the embedding classroom assessment environment less constraining and more conducive to assessment for learning than at present. I would include amongst my more salient efforts in these directions my participation in:

- The Assessment Issue Group: A change in the subject coordinator some time after the end of the study's fieldwork phase brought with it greater teacher involvement in the running of our department. Seizing the opportunity, I coordinated an assessment issue group within the department with the intention of getting teachers to collegially discuss assessment-related topics and to partake in the decision-making process. During our meetings, which were usually attended by roughly five teachers from amongst the younger ranks, it soon became apparent that they were primarily interested in discussing the computation of the Assessment Marks and the

setting of the end-of-first-year examination papers. Seeing this as a good, non-threatening starting point, I gladly went along with their wishes. As the end of the scholastic year was approaching, we decided to focus our meetings on how we could all be part of the paper setting process. But what we considered a success – for we had at least begun to work together as a group and enjoyed doing it – proved to be our undoing. The school administration not only objected to the examination papers we had produced, but practically also ‘accused’ the subject coordinator of incompetence. Although I would have liked us to continue functioning in this discussion forum, the group members grew increasingly disillusioned and started to lose interest. Sadly, the group dismantled before we had the opportunity to address our classroom assessment practices.

- The Committee for Policy and Practice of Assessment: When a colleague from another department invited me to join this newly set up school committee – her personal initiative that was however sanctioned by the school administration – I gladly accepted. The aims of the committee were: (i) to identify the strengths and weaknesses of the school’s current assessment practices; (ii) to improve on the current assessment practices; and (iii) to make suggestions for a reformed school assessment policy. All five committee members agreed that our work would also include the reconceptualisation of the traditional manner in which assessment is still regarded at school. But this went against what the school administration had in mind. For they were simply expecting us to reformulate the school assessment regulations – in particular, the promotion criteria. Matters came to a head when the administration insisted that two members from the school board would ‘guide us’ during the meetings. This development practically brought our meetings to an end.
- The review process of the Matriculation Certificate: As part of a major review exercise, the MATSEC board invited our school (and the other sixth form colleges) to forward proposals as to how the Matriculation Certificate may be reformed. The school board subsequently invited all subject coordinators to prepare a report in consultation with their teachers. We were told that our school would eventually base its proposals on these reports. In the ensuing discussion within my department, I argued particularly in favour of creating a niche for classroom assessment within the certification system – a plea that was generally welcomed by most colleagues.

This suggestion was however not included in the school's final report for reasons that had mainly to do with perceived 'logistical difficulties' and 'fears of teacher favouritism'.

Although such change initiatives, apart from landing me in a number of 'conflicts' (see Fullan, 2004), have so far failed to deliver any tangible results, I am determined not to let this dampen my commitment and enthusiasm. Consequently, when all seems black, I like to remind myself of how heartening it was, when the assessment issue group was still functioning, to see some colleagues already beginning to appreciate that there is more to classroom assessment than testing and grading. My long-term plan thus remains to rope in the support of a wide range of people and to register progress over time, even if in small, seemingly insignificant initial steps (see Weeden et al., 2002).

If nothing else, I have learned as a teacher seeking change that, as Fullan (2001) warns, putting ideas into practice is a far more complex process than people realise. What I frankly did not expect though was that the school administration in general would constitute one of the major stumbling blocks to my designs. Although it may well be that, as Fullan (2001) argues, change is only one of the forces competing for the attention of people in administrative positions, and usually not the most compelling one, I somewhat suspect that their resistance to change knows of deeper motives. It has to be appreciated that as "All innovations worth their salt call on people to question and, in some respects, to change their behaviors and their beliefs" (Fullan, 2004, p. 49), people 'caught' in change feel anxious, fearful, confused, overwhelmed, deskilled, cautious, and – if they have moral purpose – deeply disturbed (Fullan, 2004). Given this scenario, I can imagine how 'bad' my school administrators probably feel about attempts to introduce change that are not only pushed bottom-up, but also by someone (i.e., myself) who has in the past stood up on a number of occasions to be counted, even when this contrasted openly with the views of the administration. But whilst I fully understand that their resistance to change can be instructive and should be taken seriously (see Hargreaves & Fullan, 1998; cited in Fullan, 2001), I have no intention to let it carry the day. The manner in which I am trying to deal with this unexpected, disconcerting reality is however another story ...

REFERENCES

- Adler, P. A., & Adler, P. (1994). Observational techniques. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 377-392). Thousand Oaks, CA: SAGE Publications.
- Airasian, P. W. (2000). *Assessment in the classroom: a concise approach* (2nd ed.). Boston: McGraw-Hill.
- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: how researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 62(3), 315-343.
- Anderson, L. W. (2003). *Classroom assessment: enhancing the quality of teacher decision making*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Assessment Reform Group (ARG) (1999). *Assessment for learning: beyond the black box* [Pamphlet]. Cambridge: University of Cambridge, School of Education.
- Assessment Reform Group (ARG) (2002). *Assessment for learning: 10 principles* [Leaflet/poster].
- Atkinson, P., & Coffey, A. (1997). Analysing documentary realities. In D. Silverman (Ed.), *Qualitative research: theory, method and practice* (pp. 45-62). London: SAGE Publications.
- Bartolo, E. (1997, November 23). Small is painful (Address by the Minister of Education at the International Conference on Human Resources and Future Generations in Islands and Small States, Foundation for International Studies, Valletta, 6 November 1997). *The Sunday Times* [Malta], p. 38.
- Bassey, M. (1999). *Case study research in educational settings*. Buckingham: Open University Press.
- Baszanger, I., & Dodier, N. (1997). Ethnography: relating the part to the whole. In D. Silverman (Ed.), *Qualitative research: theory, method and practice* (pp. 8-23). London: SAGE Publications.
- Bezzina, C. (1997). Encouraging motivation among all staff. In J. Muscat (Ed.), *Towards the new millennium: the changing role of the teacher* (pp. 59-66). Malta: MUT Publications Ltd.
- Bezzina, C. (2001). *On becoming an effective teacher: an introductory handbook*. Malta: Indigobooks.
- Bezzina, C., Bezzina, N. R., & Stanyer, R. (2004). Exploring beginning teachers' perceptions of their preparation and professional development in Malta. *Mediterranean Journal of Educational Studies*, 9(2), 39-70.

- Bezzina, C., & Camilleri, A. (2001). The professional development of teachers in Malta. *European Journal of Education, 14*(2), 157-170.
- Biggs, J. (1996). Assessing learning quality: reconciling institutional, staff and educational demands. *Assessment and Evaluation in Higher Education, 21*(1), 5-15.
- Bishop, A. J. (1991). *Mathematical enculturation: a cultural perspective on mathematics education*. Dordrecht: Kluwer Academic Publishers.
- Black, P. (1998). *Testing: friend or foe? The theory and practice of assessment and testing*. London: The Falmer Press.
- Black, P. (1999). Assessment, learning theories and testing systems. In P. Murphy (Ed.), *Learners, learning and assessment* (pp. 118-134). London: Paul Chapman Publishing (in association with The Open University).
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: putting it into practice*. Maidenhead: Open University Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice, 5*(1), 7-74.
- Boaler, J. (1997). *Experiencing school mathematics: teaching styles, sex and setting*. Buckingham: Open University Press.
- Boaler, J. (1998). Open and closed mathematics: student experiences and understandings. *Journal for Research in Mathematics Education, 29*(1), 41-62.
- Boaler, J., Wiliam, D., & Brown, M. (1998). Students' experiences of ability grouping – disaffection, polarisation and the construction of failure. In P. Gates (Ed.), *Proceedings: first International Mathematics Education and Society Conference* (pp. 367-382). Nottingham: Centre for the Study of Mathematics Education, University of Nottingham.
- Bogdan, R. C., & Biklen, S. K. (1982). *Qualitative research for education: an introduction to theory and methods*. Boston: Allyn & Bacon.
- Borg, C., Camilleri, J., Mayo, P., & Xerri, T. (1995). Malta's national curriculum: a critical analysis. *International Review of Education, 41*(5), 337-356.
- Borg, M. G., & Falzon, J. M. (1995). Birthdate and sex effects on the scholastic attainment of primary schoolchildren: a cross-sectional study. *British Educational Research Journal, 21*(1), 61-74.
- Borg, M. G., Falzon, J. M., & Sammut, A. (1995). Age and sex differences in performance in an 11-plus selective examination. *Educational Psychology, 15*(4), 433-443.
- Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development*. Thousand Oaks, CA: SAGE Publications.

- Bright, G. W., & Joyner, J. M. (1998). Understanding and improving classroom assessment: summary of issues raised. In G. W. Bright & J. M. Joyner (Eds.), *Classroom assessment in mathematics: views from a National Science Foundation working conference* (pp. 27-57). Lanham, MD: University Press of America.
- Briscoe, G. (1998, October 11). MATSEC matters. *The Sunday Times* [Malta], p. 55.
- Broadfoot, P. (1995). Performance assessment in perspective: international trends and current English experience. In H. Torrance (Ed.), *Evaluating authentic assessment: problems and possibilities in new approaches to assessment* (pp. 9-43). Buckingham: Open University Press.
- Broadfoot, P. M. (1996). *Education, assessment and society: a sociological analysis*. Buckingham: Open University Press.
- Broadfoot, P. (2000). Preface. In A. Filer (Ed.), *Assessment: social practice and social product* (pp. ix-xii). London: RoutledgeFalmer.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of 'Assessment in Education'. *Assessment in Education: Principles, Policy and Practice*, 11(1), 7-27.
- Brookhart, S. M. (1999). *The art and science of classroom assessment: the missing part of the pedagogy*. Washington, DC: The George Washington University.
- Brown, A. L., & Ferrara, R. (1985). Diagnosing zones of proximal development. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 273-305). Cambridge: Cambridge University Press.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Brown, M. (1989). Graded assessment projects: similarities and differences. In P. Murphy & B. Moon (Eds.), *Developments in learning and assessment* (pp. 300-311). London: Hodder & Stoughton.
- Brown, S. (1990). Assessment: a changing practice. In T. Horton (Ed.), *Assessment debates* (pp. 5-11). London: Hodder & Stoughton.
- Bryant, D., & Driscoll, M. (1998). *Exploring classroom assessment in mathematics: a guide for professional development*. Reston, VA: National Council of Teachers of Mathematics.
- Buhagiar, M. A. (2003). The setting up of the University of Malta Junior College: origins, motives and polemics. *Journal of Maltese Education Research*, 1(1), 143-166. [On-line]. Available: <http://www.educ.um.edu.mt/jmer>.
- Buhagiar, M. A. (2004). 'How appropriate is this task for my class?' Exploring teachers' classroom decision-making processes as they waver between 'practical' and 'ideal' positions. *Mediterranean Journal of Educational Studies*, 9(2), 83-108.

- Burgess, R. G. (1982a). The unstructured interview as a conversation. In R. G. Burgess (Ed.), *Field research: a sourcebook and field manual* (pp. 107-110). London: Routledge.
- Burgess, R. G. (1982b). Some role problems in field research. In R. G. Burgess (Ed.), *Field research: a sourcebook and field manual* (pp. 45-49). London: Routledge.
- Burgess, R. G. (1984). *In the field: an introduction to field research*. London: Routledge.
- Calderhead, J. (1987). Introduction. In J. Calderhead (Ed.), *Exploring teachers' thinking* (pp. 1-19). London: Cassell.
- Calderhead, J. (1988). Introduction. In J. Calderhead (Ed.), *Teachers' professional learning* (pp. 1-11). London: The Falmer Press.
- Calfee, R. C., & Masuda, W. V. (1997). Classroom assessment as inquiry. In G. D. Phye (Ed.), *Handbook of classroom assessment: learning, adjustment, and achievement* (pp. 69-102). San Diego, CA: Academic Press.
- Calleja, J. (1988). Thoughts on the concepts and practices of the Maltese educational system. In C. J. Farrugia (Ed.), *Education in Malta: a look to the future* (pp. 27-38). Malta: UNESCO.
- Camilleri, A. (1995). *Bilingualism in education: the Maltese experience*. Heidelberg: Julius Groos Verlag.
- Camilleri, P. (1995, June 4). Why a Junior College. *The Sunday Times* [Malta], p. 6.
- Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 19-32). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Carter, K., & Doyle, W. (1987). Teachers' knowledge structures and the comprehension processes. In J. Calderhead (Ed.), *Exploring teachers' thinking* (pp. 147-160). London: Cassell.
- Central Office of Statistics (1997). *Education statistics 1995-1996*. Malta: Department of Information.
- César, M. (1998). Social interactions and mathematics learning. In P. Gates (Ed.), *Proceedings: first International Mathematics Education and Society Conference* (pp. 110-119). Nottingham: Centre for the Study of Mathematics Education, University of Nottingham.
- Chalmers, R., Thake, A. M., & Sciberras, J. (2004). *State higher education funding* (A Report of the State Higher Education Funding Working Group to the Minister of Education, Youth and Employment). Malta: Ministry of Education, Youth and Employment.

- Chetcuti, D. (1998). *The physics secondary education certificate examination: a Maltese case study*. Unpublished PhD thesis, Nottingham Trent University, UK.
- Chetcuti, D., & Griffiths, M. (2002). The implications for student self-esteem of ordinary differences in schools: the cases of Malta and England. *British Educational Research Journal*, 28(4), 529-549.
- Chircop, D. (1997). Voting with their feet: students and absenteeism. In R. G. Sultana (Ed.), *Inside/outside schools: towards a critical sociology of education in Malta* (pp. 353-374). Malta: Publishers Enterprises Group (PEG) Ltd.
- Cizek, G. J. (1997). Learning, achievement, and assessment: constructs at a crossroads. In G. D. Phye (Ed.), *Handbook of classroom assessment: learning, adjustment, and achievement* (pp. 1-32). San Diego, CA: Academic Press.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1995). Teachers' assessment practices: preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159-179.
- Cole, A. L. (1997). Impediments to reflective practice: toward a new agenda for research on teaching. *Teachers and Teaching: Theory and Practice*, 3(1), 7-27.
- Coleiro, L. (2004). *OHS survey carried out at the G. F. Abela Junior College*. [On-line]. Available: <http://jc.um.edu.mt/admin/jcohs.pdf>.
- Creswell, J. W. (1994). *Research design: qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE Publications.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.
- Cross, K. P. (1998). Classroom research: implementing the scholarship of teaching. In T. Angelo (Ed.), *Classroom assessment and research: an update on uses, approaches, and research findings* (pp. 5-12). San Francisco: Jossey-Bass Publishers.
- Cumming, J. J., & Maxwell, G. S. (1999). Contextualising authentic assessment. *Assessment in Education: Principles, Policy and Practice*, 6(2), 177-194.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review*, 64(1), 5-30.
- Darmanin, M. (1992). The labour market of schooling: Maltese girls in education and economic planning. *Gender and Education*, 4(1-2), 105-126.
- Day, C. (1993). The importance of learning biography in supporting teacher development: an empirical study. In C. Day, J. Calderhead & P. Denicolo (Eds.), *Research on teacher thinking: understanding professional development* (pp. 221-232). London: The Falmer Press.

- Day, C. (1999). *Developing teachers: the challenges of lifelong learning*. London: The Falmer Press.
- Day, C. (2000). Stories of change and professional development: the cost of commitment. In C. Day, A. Fernandez, T. E. Hauge & J. Møller (Eds.), *The life and works of teachers: international perspectives in changing times* (pp. 109-129). London: The Falmer Press.
- Day, T. (2005). Teachers' craft knowledge: a constant in times of change? *Irish Educational Studies*, 24(1), 21-30.
- Delandshere, G. (2001). Implicit theories, unexamined assumptions, and the status quo of educational assessment. *Assessment in Education: Principles, Policy and Practice*, 8(2), 113-133.
- de Lange, J. (1995). Assessment: no change without problems. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment* (pp. 87-172). Albany, NY: State University of New York Press.
- Delap, M. R. (1994). An investigation into the accuracy of A-Level predicted grades. *Educational Research*, 36(2), 135-148.
- Delap, M. R. (1995). Teachers' estimates of candidates' performance in public examinations. *Assessment in Education: Principles, Policy and Practice*, 2(1), 75-92.
- Denvir, B. (1989). Assessment purposes and learning in mathematics education. In P. Murphy & B. Moon (Eds.), *Developments in learning and assessment* (pp. 277-289). London: Hodder & Stoughton.
- Denzin, N. K. (1994). The art and politics of interpretation. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 500-515). Thousand Oaks, CA: SAGE Publications.
- Denzin, N. K., & Lincoln, Y. S. (1994). Introduction: entering the field of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 1-17). Thousand Oaks, CA: SAGE Publications.
- Desforges, C. (1989). *Testing and assessment*. London: Cassell.
- Desforges, C., & Cockburn, A. (1987). *Understanding the mathematics teacher: a study of practice in first schools*. London: The Falmer Press.
- Doyle, W. (1988). Work in mathematics classes: the context of students' thinking during instruction. *Educational Psychologist*, 23(2), 167-180.
- Eggleston, J. (1991). Teaching teachers to assess. *European Journal of Education*, 26(3), 231-237.

- Eisner, E. W. (1993). Reshaping assessment in education: some criteria in search of practice. *Journal of Curriculum Studies*, 25(3), 219-233.
- Elbaz, F. (1990). Knowledge and discourse: the evolution of research on teacher thinking. In C. Day, M. Pope & P. Denicolo (Eds.), *Insight into teachers' thinking and practice* (pp. 15-42). London: The Falmer Press.
- Elliott, J. (1993a). Introduction. In J. Elliott (Ed.), *Reconstructing teacher education: teacher development* (pp. 1-12). London: The Falmer Press.
- Elliott, J. (1993b). Three perspectives on coherence and continuity in teacher education. In J. Elliott (Ed.), *Reconstructing teacher education: teacher development* (pp. 15-19). London: The Falmer Press.
- Ellis, A. K. (2001). *Teaching, learning and assessment together: the reflective classroom*. Larchmont, NY: Eye On Education.
- Elwood, J., & Klenowski, V. (2002). Creating communities of shared practice: the challenges of assessment use in learning and teaching. *Assessment and Evaluation in Higher Education*, 27(3), 243-256.
- Ely, M. (with Anzul, M., Friedman, T., Gardner, D., & McCormack Steinmetz, A.). (1991). *Doing qualitative research: circles within circles*. London: The Falmer Press.
- Ernest, P. (1998). A postmodern perspective on research in mathematics education. In A. Sierpiska & J. Kilpatrick (Eds.), *Mathematics education as a research domain: a search for identity* (pp. 71-85). Dordrecht: Kluwer Academic Publishers.
- European Commission (2005a). *Key facts and figures about the European Union – new member states and candidate countries – education*. [On-line]. Available: http://europa.eu.int/abc/keyfigures/candidates_member/education/index_accessible_en.htm.
- European Commission (2005b). *Structural indicators: update of the statistical annex (annex 1) to the 2005 report from the Commission to the Spring European Council*. [On-line]. Available: http://europa.eu.int/growthandjobs/pdf/statistical_annex_2005_en.pdf.
- Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research*, 38(1), 47-65.
- Farrugia, C. (1994). Malta: educational development in a small island state. In C. Farrugia (Ed.), *A new vision for primary schools* (pp. 265-288). Malta: MUT Publications Ltd.
- Farrugia, C. J., & Attard, P. A. (1989). *The multi-functional administrator: educational development in the small states of the Commonwealth*. London: The Commonwealth Secretariat.

- Fenech, J. (1988). Curriculum content: inert or dynamic knowledge? In C. J. Farrugia (Ed.), *Education in Malta: a look to the future* (pp. 71-79). Malta: UNESCO.
- Filer, A. (2000). Classroom contexts of assessment: editor's introduction. In A. Filer (Ed.), *Assessment: social practice and social product* (pp. 83-86). London: RoutledgeFalmer.
- Filer, A., & Pollard, A. (2000). Assessment and parents' strategic action. In A. Filer (Ed.), *Assessment: social practice and social product* (pp. 133-150). London: RoutledgeFalmer.
- Fontana, A., & Frey, J. H. (1994). Interviewing: the art of science. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 361-376). Thousands Oaks, CA: SAGE Publications.
- Foster, P. (1989). Change and adjustment in a further education college. In R. G. Burgess (Ed.), *The ethics of educational research* (pp. 188-204). New York: The Falmer Press.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.
- Fullan, M. (2004). *Leading in a culture of change: personal action guide and workbook*. San Francisco: Jossey-Bass Publishers.
- Gans, H. J. (1982). The participant observer as a human being: observations on the personal aspects of research. In R. G. Burgess (Ed.), *Field research: a sourcebook and field manual* (pp. 53-61). London: Routledge.
- Gates, P. (2002). Issues of equity in mathematics education: defining the problem, seeking solutions. In L. Haggarty (Ed.), *Teaching mathematics in secondary schools: a reader* (pp. 211-228). London: RoutledgeFalmer.
- G. F. Abela Junior College Regulations, The (1995). (Legal Notice 123 of 1995).
- Gipps, C. (1993). The profession of educational research. *British Educational Research Journal*, 19(1), 3-16.
- Gipps, C. V. (1994). *Beyond testing: towards a theory of educational assessment*. London: RoutledgeFalmer.
- Gipps, C. (1996). Introduction. In P. F. Murphy & C. V. Gipps (Eds.), *Equity in the classroom: towards effective pedagogy for girls and boys* (pp. 1-6). London: The Falmer Press.
- Gipps, C., Brown, M., McCallum, B., & McAlister, S. (1995). *Intuition or evidence?* Buckingham: Open University Press.
- Gipps, C., & Murphy, P. (1994). *A fair test? Assessment, achievement and equity*. Buckingham: Open University Press.

- Gipps, C., & Stobart, G. (1993). *Assessment: a teachers' guide to the issues* (2nd ed.). London: Hodder & Stoughton.
- Glaser, R. (1990). Toward new models for assessment. *International Journal of Educational Research*, 14(5), 475-483.
- Goldin, G. A. (1992). Toward an assessment framework for school mathematics. In R. Lesh & S. J. Lamon (Eds.), *Assessment of authentic performance in school mathematics* (pp. 63-88). Washington, DC: American Association for the Advancement of Science Press.
- Griffiths, M. (1998). *Educational research for social justice: getting off the fence*. Buckingham: Open University Press.
- Grima, G., & Chetcuti, D. (2003). Current assessment practices in schools in Malta and Gozo: a research report. *Journal of Maltese Education Research*, 1(2), 57-94. [On-line]. Available: <http://www.educ.um.edu.mt/jmer>.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: SAGE Publications.
- Hall, K., Webber, B., Varley, S., Young, V., & Dorman, P. (1997). A study of Teacher Assessment at Key Stage 1. *Cambridge Journal of Education*, 27(1), 107-122.
- Hammersley, M. (1992). On feminist methodology. *Sociology*, 26(2), 187-206.
- Hammersley, M. (1995). Opening up the quantitative-qualitative divide. *Education Section Review*, 19(1), 2-9.
- Hammersley, M. (2003). Can and should educational research be educative? *Oxford Review of Education*, 29(1), 3-25.
- Hammersley, M., & Atkinson, P. (1995). *Ethnography: principles in practice* (2nd ed.). London: Routledge.
- Hanafin, J. (1995). Moving beyond the figures: using quantitative methods in educational research. *Irish Educational Studies*, 14, 184-201.
- Hanson, F. A. (2000). How tests create what they are intended to measure. In A. Filer (Ed.), *Assessment: social practice and social product* (pp. 67-81). London: RoutledgeFalmer.
- Hargreaves, A. (1984). The significance of classroom coping strategies. In A. Hargreaves & P. Woods (Eds.), *Classrooms and staffrooms: the sociology of teachers and teaching* (pp. 64-85). Milton Keynes: Open University Press.
- Hargreaves, A. (1994). *Changing teachers, changing times: teachers' work and culture in the postmodern age*. London: Continuum.

- Hargreaves, D. H. (1993). A common-sense model of the professional development of teachers. In J. Elliott (Ed.), *Reconstructing teacher education: teacher development* (pp. 86-92). London: The Falmer Press.
- Harlen, W. (1994a). Issues and approaches to quality assurance and quality control in assessment. In W. Harlen (Ed.), *Enhancing quality in assessment* (pp. 11-25). London: Paul Chapman Publishing Ltd.
- Harlen, W. (1994b). Towards quality in assessment. In W. Harlen (Ed.), *Enhancing quality in assessment* (pp. 139-145). London: Paul Chapman Publishing Ltd.
- Harlen, W., Gipps, C., Broadfoot, P., & Nuttall, D. (1992). Assessment and the improvement of education. *The Curriculum Journal*, 3(3), 215-230.
- Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy and Practice*, 4(3), 365-379.
- Harris, A. (1998). Effective teaching: a review of the literature. *School Leadership and Management*, 18(2), 169-183.
- Hildebrand, G. M. (1996). Redefining achievement. In P. F. Murphy & C. V. Gipps (Eds.), *Equity in the classroom: towards effective pedagogy for girls and boys* (pp. 149-172). London: The Falmer Press.
- Hockey, J. (1993). Research methods – researching peers and familiar settings. *Research Papers in Education*, 8(2), 199-225.
- Hodder, I. (1994). The interpretation of documents and material culture. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 393-402). Thousand Oaks, CA: SAGE Publications.
- Hollis, M. (1994). *The philosophy of social science: an introduction*. Cambridge: Cambridge University Press.
- Hopkins, D. (1994). School improvement in an ERA of change. In P. Ribbins & E. Burrige (Eds.), *Improving education: promoting quality in schools* (pp. 74-91). London: Cassell.
- Howe, K. R. (1998). The interpretive turn and the new debate in education. *Educational Researcher*, 27(8), 13-20.
- Hughes, J., & Sharrock, W. (1997). *The philosophy of social research* (3rd ed.). London: Longman.
- Impara, J. C., Plake, B. S., & Fager, J. J. (1993). Teachers' assessment background and attitude toward testing. *Theory into Practice*, 32(2), 113-117.
- Ingersoll, R. M. (1996). Teachers' decision-making power and school conflict. *Sociology of Education*, 69(2), 159-176.

- Jaworski, B. (2002). Social constructivism in mathematics learning and teaching. In L. Haggarty (Ed.), *Teaching mathematics in secondary schools: a reader* (pp. 67-82). London: RoutledgeFalmer.
- Jaworski, B. (2004). Grappling with complexity: co-learning in inquiry communities in mathematics teaching development. In M. Johnsen Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (pp. 17-36). Bergen: Bergen University College.
- Joyner, J. M. (1998). Thoughts on classroom assessment. In G. W. Bright & J. M. Joyner (Eds.), *Classroom assessment in mathematics: views from a National Science Foundation working conference* (pp. 185-189). Lanham, MD: University Press of America.
- Junior College (1998). *Look ... before you leap*. Malta: Author.
- Junior College (2000a). *Junior College strategic plan 2000-2005*. Unpublished.
- Junior College (2000b). *The University of Malta Junior College – a handbook* (5th ed.). Malta: Author.
- Kvale, S. (1996). *Interviews: an introduction to qualitative research interviewing*. Thousand Oaks, CA: SAGE Publications.
- LaCelle-Peterson, M. (2000). How assessment policies and practices obscure the education of language minority students. In A. Filer (Ed.), *Assessment: social practice and social product* (pp. 27-42). London: RoutledgeFalmer.
- Lesh, R., Lamon, S. J., Behr, M., & Lester, F. (1992). Future directions for mathematics assessment. In R. Lesh & S. J. Lamon (Eds.), *Assessment of authentic performance in school mathematics* (pp. 379-425). Washington, DC: American Association for the Advancement of Science Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: SAGE Publications.
- Lincoln, Y. S., & Guba, E. G. (1990). Judging the quality of case study reports. *International Journal of Qualitative Studies in Education*, 3(1), 53-59.
- Loyd, B. H., & Loyd, D. E. (1997). Kindergarten through grade 12 standards: a philosophy of grading. In G. D. Phye (Ed.), *Handbook of classroom assessment: learning, adjustment, and achievement* (pp. 481-489). San Diego, CA: Academic Press.
- Mallia, A., Baldacchino, G., Mayo, P., Wain, K., Buhagiar, A., & Macelli, N. (2001). *Malta: report on the national consultation process on lifelong learning: outcomes to date and plan for the way forward*. [On-line]. Available: http://europa.eu.int/comm/education/policies/III/life/report/candidate/malta_en.pdf.

- Malta Union of Teachers (MUT) (1996). *Mozzjoni 4: Eżamijiet – xkiel jew motivazzjoni?* [Motion 4: Exams – hurdle or motivation?]. Motion approved by MUT delegates during the Annual General Conference, 3-5 June 1996.
- Malta Union of Teachers (MUT) (1998). Memorandum to political parties. *MUT Newsletter*, 4/98, v-vii.
- Mansueto, R. (1997). Examinations and stress: childhood and education subverted. In R. G. Sultana (Ed.), *Inside/outside schools: towards a critical sociology of education in Malta* (pp. 183-200). Malta: Publishers Enterprises Group (PEG) Ltd.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: 1. Outcome and process. *British Journal of Educational Psychology*, 46, 4-11.
- Mathematical Sciences Education Board (MSEB) (1993). *Measuring what counts: a conceptual guide for mathematics assessment*. Washington, DC: National Academy Press.
- Matriculation Certificate Examination Regulations, The (1995). (Legal Notice 112 of 1995).
- MATSEC (1994a). *The Matriculation Certificate Examination (Intermediate Level): 1996-1998 regulations and syllabuses*. Malta: Author.
- MATSEC (1994b). *The Matriculation Certificate Examination to be offered as from May 1996* (promotional pamphlet).
- Mavrommatis, Y. (1997). Understanding assessment in the classroom: phases of the assessment process – the assessment episode. *Assessment in Education: Principles, Policy and Practice*, 4(3), 381-399.
- Maykut, P., & Morehouse, R. (1994). *Beginning qualitative research: a philosophical and practical guide*. London: The Falmer Press.
- McCallum, B., McAlister, S., Brown, M., & Gipps, C. (1993). Teacher assessment at key stage one. *Research Papers in Education*, 8(3), 305-327.
- McNamara, D. R. (1980). The outsider's arrogance: the failure of participant observers to understand classroom events. *British Educational Research Journal*, 6(2), 113-125.
- Mercieca, H. (1997). A defence of streaming. *The Teacher*, 66, 6-7.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education* (2nd ed.). San Francisco: Jossey-Bass Publishers.
- Mifsud, J. (1991). Assessment: from exam orientation to classroom practice. In R. G. Sultana (Ed.), *Themes in education: a Maltese reader* (pp. 113-130). Malta: Mireva Publications.

- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook* (2nd ed.). Thousand Oaks, CA: SAGE Publications.
- Ministry of Education (1999). *Creating the future together: national minimum curriculum*. Malta: Author.
- Ministry of Education (2001). *Education in Malta – a handbook*. [On-line]. Available: http://www.education.gov.mt/edu/educ_handbook.htm.
- Ministry of Education, Youth and Employment (2005). *For all children to succeed: a new network organisation for quality education in Malta*. Malta: Author.
- Mitchell, I. (1999). Bridging the gulf between research and practice. In J. Loughran (Ed.), *Researching teaching: methodologies and practices for understanding pedagogy* (pp. 44-64). London: The Falmer Press.
- Moody, I. (2001). A case-study of the predictive validity and reliability of Key Stage 2 test results, and teacher assessments, as baseline data for target-setting and value-added at Key Stage 3. *The Curriculum Journal*, 12(1), 81-101.
- Morse, J. M. (1994). Designing funded qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220-235). Thousands Oaks, CA: SAGE Publications.
- Murphy, P. (1996). Defining pedagogy. In P. F. Murphy & C. V. Gipps (Eds.), *Equity in the classroom: towards effective pedagogy for girls and boys* (pp. 9-22). London: The Falmer Press.
- Murphy, R., & Torrance, H. (1988). *The changing face of educational assessment*. Milton Keynes: Open University Press.
- Murphy, R., & Torrance, H. (1990). The need for change. In T. Horton (Ed.), *Assessment debates* (pp. 12-18). London: Hodder & Stoughton.
- National Council of Teachers of Mathematics (NCTM) (1995). *Assessment standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Statistics Office (2001). *Education statistics 1999-2000*. Malta: Department of Information.
- Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.
- Nias, J., Southworth, G., & Yeomans, R. (1989). *Staff relations in the primary school: a study of organizational cultures*. London: Cassell.

- Nuttall, D. L. (1989). The validity of assessments. In P. Murphy & B. Moon (Eds.), *Developments in learning and assessment* (pp. 265-276). London: Hodder & Stoughton.
- Nuttall, D. L. (1993). Presentation at Centre for Policy Studies Conference, 21 September 1993. Paper 20, In R. Murphy & P. Broadfoot (Eds.) (1995), *Effective assessment and the improvement of education: a tribute to Desmond Nuttall* (pp. 236-241). London: The Falmer Press.
- Office of Standards in Education (OFSTED) (1998). *Secondary education 1993-97: a review of secondary schools in England*. London: HMSO.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: SAGE Publications.
- Pennycook, D. (1991). Moderation of continuous assessment systems in developing countries. *Compare*, 21(2), 145-152.
- Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes: towards a wider conceptual field. *Assessment in Education: Principles, Policy and Practice*, 5(1), 85-102.
- Peshkin, A. (1988a). In search of subjectivity – one's own. *Educational Researcher*, 17(7), 17-21.
- Peshkin, A. (1988b). Understanding complexity: a gift of qualitative inquiry. *Anthropology and Education Quarterly*, 19(4), 416-424.
- Peshkin, A. (1993). The goodness of qualitative research. *Educational Researcher*, 22(2), 23-29.
- Peterson, J., & Stack, C. (1998). A Minnesota story: a system approach to Classroom Assessment and Research. In T. Angelo (Ed.), *Classroom Assessment and Research: an update on uses, approaches, and research findings* (pp. 67-77). San Francisco: Jossey-Bass Publishers.
- Phye, G. D. (1997a). Classroom assessment: a multidimensional perspective. In G. D. Phye (Ed.), *Handbook of classroom assessment: learning, adjustment, and achievement* (pp. 33-51). San Diego, CA: Academic Press.
- Phye, G. D. (1997b). Epilogue: classroom assessment – looking forward. In G. D. Phye (Ed.), *Handbook of classroom assessment: learning, adjustment, and achievement* (pp. 531-539). San Diego, CA: Academic Press.
- Plake, B. S., & Impara, J. C. (1997). Teacher assessment literacy: what do teachers know about assessment? In G. D. Phye (Ed.), *Handbook of classroom assessment: learning, adjustment, and achievement* (pp. 53-68). San Diego, CA: Academic Press.

- Pollard, A. (1985). Opportunities and difficulties of a teacher-ethnographer: a personal account. In R. G. Burgess (Ed.), *Field methods in the study of education* (pp. 217-233). London: The Falmer Press.
- Pollard, A. (1998). [Review of the book *Inside/outside schools: towards a critical sociology of education in Malta*]. *Mediterranean Journal of Educational Studies*, 3(2), 174-178.
- Pollard, A. (2002). *Reflective teaching: effective and evidence informed professional practice*. London: Continuum.
- Popham, W. J. (1991). Appropriateness of teachers' test-preparation practices. *Educational Measurement: Issues and Practices*, 10(4), 12-15.
- Popham, W. J. (1999). *Classroom assessment: what teachers need to know* (2nd ed.). Boston: Allyn & Bacon.
- Post-Secondary Level NMC (1991). (Legal Notice 56 of 1991).
- Pound, T. (1998). Forty years on: the issue of breadth in the post-16 curriculum. *Oxford Review of Education*, 24(2), 167-180.
- Pring, R. (1984). The problems of confidentiality. In M. Skilbeck (Ed.), *Evaluating the curriculum of the eighties* (pp. 38-44). London: Hodder & Stoughton.
- Pryor, J., & Torrance, H. (2000). Questioning the three bears: the social construction of classroom assessment. In A. Filer (Ed.), *Assessment: social practice and social product* (pp. 110-128). London: RoutledgeFalmer.
- Punch, M. (1994). Politics and ethics in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 83-97). Thousands Oaks, CA: SAGE Publications.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, 4-13.
- Romberg, T. A., & Kaput, J. J. (1999). Mathematics worth teaching, mathematics worth understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 3-17). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Romberg, T. A., & Wilson, L. D. (1995). Issues related to the development of an authentic assessment system for school mathematics. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment* (pp. 1-18). Albany, NY: State University of New York Press.
- Rowntree, D. (1987). *Assessing students: how shall we know them?* (2nd ed.). London: Kogan Page.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.

- Sammut, A. (1994). Completing primary schooling: an analysis of results. In C. Farrugia (Ed.), *A new vision for primary schools* (pp. 77-88). Malta: MUT Publications Ltd.
- Schafer, W. D. (1993). Assessment literacy for teachers. *Theory into Practice*, 32(2), 118-126.
- Schoenfeld, A. H. (1988). When good teaching leads to bad results: the disasters of 'well-taught' mathematics courses. *Educational Psychologist*, 23(2), 145-166.
- Schoenfeld, A. H. (1999). Looking toward the 21st century: challenges to educational theory and practice. *Educational Researcher*, 28(7), 4-14.
- Schoenfeld, A. H. (2002). Making mathematics work for all children: issues of standards, testing, and equity. *Educational Researcher*, 31(1), 13-25.
- Schön, D. (1983). *The reflective practitioner: how professionals think in action*. New York: Basic Books.
- Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118-137). Thousand Oaks, CA: SAGE Publications.
- Secondary Education Certificate Examination Regulations, The (1995). (Legal Notice 111 of 1995).
- Seegers, G., & Gravemeijer, K. (1997). Implementation and effect of realistic curricula. In M. Beishuizen, K. P. E. Gravemeijer & E. C. D. M. van Lieshout (Eds.), *The role of contexts and models in the development of mathematical strategies and procedures* (pp. 255-272). Utrecht: CD β Press.
- Seidman, I. E. (1998). *Interviewing as qualitative research: a guide for researchers in education and social sciences* (2nd ed.). New York: Teachers College Press.
- Shafer, M. C., & Romberg, T. A. (1999). Assessment in classrooms that promote understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 159-184). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Shepard, L. A. (1991). Psychometricians' beliefs about learning. *Educational Researcher*, 20(6), 2-16.
- Shepard, L. A. (1995). Using assessment to improve learning. *Educational Leadership*, 52(5), 38-43.
- Silverman, D. (1993). *Interpreting qualitative data: methods for analysing talk, text and interaction*. London: SAGE Publications.

- Simons, H. (1982). Suggestions for a school self-evaluation based on democratic principles. In R. McCormick, J. Bynner, P. Clift, M. James & C. Morrow Brown (Eds.), *Calling education to account* (pp. 286-295). London: Heinemann (in association with The Open University).
- Simons, H. (1989). Ethics of case study in educational research and evaluation. In R. G. Burgess (Ed.), *The ethics of educational research* (pp. 114-138). New York: The Falmer Press.
- Smith, M. L. (1991). Put to the test: the effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.
- Southworth, G. (1994). The learning school. In P. Ribbins & E. Burrige (Eds.), *Improving education: promoting quality in schools* (pp. 52-73). London: Cassell.
- Sowder, J. T. (1998). Ethics in mathematics education research. In A. Sierpiska & J. Kilpatrick (Eds.), *Mathematics education as a research domain: a search for identity* (pp. 427-442). Dordrecht: Kluwer Academic Publishers.
- Stake, R. E. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236-247). Thousands Oaks, CA: SAGE Publications.
- Steadman, M. (1998). Using Classroom Assessment to change both teaching and learning. In T. Angelo (Ed.), *Classroom Assessment and Research: an update on uses, approaches, and research findings* (pp. 23-35). San Francisco: Jossey-Bass Publishers.
- Steadman, M., & Svinicki, M. (1998). CATs: a student's gateway to better learning. In T. Angelo (Ed.), *Classroom Assessment and Research: an update on uses, approaches, and research findings* (pp. 13-20). San Francisco: Jossey-Bass Publishers.
- Stefani, L. A. J. (1998). Assessment in partnership with learners. *Assessment and Evaluation in Higher Education*, 23(4), 339-350.
- Stiggins, R. (1988). Revitalizing classroom assessment: the highest instructional priority. *Phi Delta Kappan*, 69, 363-368.
- Stiggins, R. J. (1992). High quality classroom assessment: what does it really mean? *Educational Measurement: Issues and Practice*, 11(2), 35-39.
- Stiggins, R. J. (1997). *Student-centered classroom assessment* (2nd ed.). Upper Saddle River, NJ: Merrill.
- Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement*, 22(4), 271-286.
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: investigating the practices of classroom assessment*. Albany, NY: State University of New York Press.

- Stiggins, R. J., Griswold, M. M., & Wikelund, K. R. (1989). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement*, 26(3), 233-246.
- Sultana, R. G. (1995). Vocational secondary schools in Malta: quality of education and the reproduction of inequality. *The Vocational Aspect of Education*, 47(1), 51-67.
- Sultana, R. G. (1997). Educational development in post-colonial Malta. In R. G. Sultana (Ed.), *Inside/outside schools: towards a critical sociology of education in Malta* (pp. 87-117). Malta: Publishers Enterprises Group (PEG) Ltd.
- Sultana, R. G. (1998). Malta. In M. Bray & L. Steward (Eds.), *Examination systems in small states: comparative perspectives on policies, models and operations*. London: Commonwealth Secretariat.
- Swan, M. (2001). Dealing with misconceptions in mathematics. In P. Gates (Ed.), *Issues in mathematics teaching* (pp. 147-165). London: RoutledgeFalmer.
- Task Group on Assessment and Testing (TGAT) (1988). *A report*. London: Department of Education and Science.
- Tolley, G. (1989). Learning and assessment. In P. Murphy & B. Moon (Eds.), *Developments in learning and assessment* (pp. 254-259). London: Hodder & Stoughton.
- Torrance, H. (1995). The role of assessment in educational reform. In H. Torrance (Ed.), *Evaluating authentic assessment: problems and possibilities in new approaches to assessment* (pp. 144-156). Buckingham: Open University Press.
- Torrance, H. (2000). Postmodernism and educational assessment. In A. Filer (Ed.), *Assessment: social practice and social product* (pp. 173-188). London: RoutledgeFalmer.
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: teaching, learning and assessment in the classroom*. Buckingham: Open University Press.
- Ventura, F., & Murphy, R. (1998). The impact of measures to promote equity in the Secondary Education Certificate in Malta: an evaluation. *Mediterranean Journal of Educational Studies*, 3(1), 47-73.
- von Glasersfeld, E. (1989). Learning as a constructive activity. In P. Murphy & B. Moon (Eds.), *Developments in learning and assessment* (pp. 5-18). London: Hodder & Stoughton.
- von Glasersfeld, E. (1995). A constructivist approach to teaching. In L. P. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 3-15). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.

- Wain, K. (1991). *The Maltese National Curriculum: a critical evaluation*. Malta: Mireva Educational.
- Wain, K. (1995, July 9). The Junior College: misconceptions, conspiracy theories and the rest. *The Sunday Times* [Malta], p. 62.
- Wain, K., Attard, P., Bezzina, C., Darmanin, M., Farrugia, C., Psaila, A., Sammut, J., Sultana, R., & Zammit, L. (1995). *Tomorrow's schools: developing effective learning cultures*. Malta: Ministry of Education and Human Resources.
- Walker, C., & Angelo, T. (1998). A collective effort Classroom Assessment Technique: promoting high performance in student teams. In T. Angelo (Ed.), *Classroom Assessment and Research: an update on uses, approaches, and research findings* (pp. 101-112). San Francisco: Jossey-Bass Publishers.
- Watson, A. (2001). Making judgements about pupils' mathematics. In P. Gates (Ed.), *Issues in mathematics teaching* (pp. 217-231). London: RoutledgeFalmer.
- Webb, N. L. (1998). Thoughts on assessment in the mathematics classroom. In G. W. Bright & J. M. Joyner (Eds.), *Classroom assessment in mathematics: views from a National Science Foundation working conference* (pp. 101-114). Lanham, MD: University Press of America.
- Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment: what's in it for schools?* London: RoutledgeFalmer.
- Weiner, G. (1994). *Feminisms in education: an introduction*. Buckingham: Open University Press.
- Whyte, W. F. (1982). Interviewing in field research. In R. G. Burgess (Ed.), *Field research: a sourcebook and field manual* (pp. 111-122). London: Routledge.
- Wiliam, D. (1996). Standards in examinations: a matter of trust? *The Curriculum Journal*, 7(3), 293-306.
- Wiliam, D. (1998). Enculturating learners into communities of practice: raising achievement through classroom assessment. Paper presented at the European Conference on Educational Research, University of Ljubljana, Slovenia, 17-20 September 1998. [On-line]. Available: <http://www.kcl.ac.uk/depsta/education/publications/ECER98.pdf>.
- Wiliam, D. (2001). An overview of the relationship between assessment and the curriculum. In D. Scott (Ed.), *Curriculum and assessment* (pp. 165-181). Westport, CT: Ablex Publishing.
- Wiliam, D., & Black, P. (1996). Meanings and consequences: a basis for distinguishing formative from summative functions of assessment? *British Educational Research Journal*, 22(5), 537-548.

- Winch, C., & Gingell, J. (1999). *Key concepts in the philosophy of education*. London: Routledge.
- Wolcott, H. F. (1990). *Writing up qualitative research*. Newbury Park, CA: SAGE Publications.
- Wolcott, H. F. (1994). *Transforming qualitative data: description, analysis, and interpretation*. Thousand Oaks, CA: SAGE Publications.
- Wood, R. (1990). The agenda for educational measurement. In T. Horton (Ed.), *Assessment debates* (pp. 48-56). London: Hodder & Stoughton.
- Woods, P. (1990). *Teacher skills and strategies*. London: The Falmer Press.
- Yin, R. K. (1994). *Case study research: design and methods* (2nd ed.). Thousand Oaks, CA: SAGE Publications.
- Zammit Mangion, J. (1988). An analysis of the growth and development of education in Malta since 1946. In C. J. Farrugia (Ed.), *Education in Malta: a look to the future* (pp. 17-25). Malta: UNESCO.
- Zammit Mangion, J. (1992). *Education in Malta*. Malta: Studia Editions.
- Zammit Mangion, J. (1994, September 15 and 16). The new University entry requirements (a two-part article). *The Times* [Malta], pp. 5-6 (September 15) and pp. 5-6, 8 (September 16).
- Zammit Mangion, J. (1995, June 18). The Junior College: some undisclosed issues. *The Sunday Times* [Malta], p. 14.
- Zarb Adami, M., Debono, S., & Sammut, F. (1999). *The MATSEC examinations*. Unpublished report by the MATSEC Analysis Facilitating Board.
- Zeni, J. (1998). A guide to ethical issues and action research. *Educational Action Research*, 6(1), 9-19.

APPENDIX I

The Contrasting Pairs of Statements and Cards

The fifteen contrasting pairs of statements used in the ‘mathematics interview’ were:

- Pair 1:** A student’s learning of mathematics depends mainly on studying hard / A student’s learning of mathematics depends mainly on intelligence
- Pair 2:** A ‘good’ mathematics teacher is born / Anyone who works hard can become a ‘good’ mathematics teacher
- Pair 3:** New mathematical knowledge is primarily the result of mathematicians’ hard work / New mathematical knowledge is primarily the result of mathematicians’ intelligence
- Pair 4:** Students learn more in class by investigating and exploring / Students learn more in class by listening and practising
- Pair 5:** Following closely teacher’s reasoning is a sign of student learning / Coming up with original ideas is a sign of student learning
- Pair 6:** A student learns more by following closely teacher’s methods / A student learns more by coming up with and trying out own methods
- Pair 7:** I prefer to use teaching materials from published sources / I prefer to create my own teaching materials
- Pair 8:** I would feel embarrassed if I could not work out a problem in class / I would not feel embarrassed if I could not work out a problem in class
- Pair 9:** It is better if a teacher first works out a problem in class and then gives students similar problems for practice / It is better if a teacher first presents students with a problem in class and then invites them to work it out
- Pair 10:** Mathematicians work like artists / Mathematicians work like explorers
- Pair 11:** Mathematicians create mathematical knowledge / Mathematicians discover mathematical knowledge
- Pair 12:** Students learn more when they work on their own / Students learn more when they work in a group
- Pair 13:** As a teacher I feel more comfortable working within a team / As a teacher I feel more comfortable working on my own
- Pair 14:** I prefer to set individual tasks in class / I prefer to set group tasks in class
- Pair 15:** Mathematicians typically work in isolation / Mathematicians typically work in groups

These contrasting pairs of statements were presented to the interviewees on separate cards. One of these cards (i.e., pair 4) is reproduced below:

Students learn more
in class by
investigating and
exploring

Students learn more
in class by
listening and
practising

APPENDIX II

The Four Tasks

Task 1

Express $\frac{2x - 8}{(x^2 + 4)(x + 1)}$ in partial fractions and hence or otherwise evaluate

$$\int_0^1 \frac{2x - 8}{(x^2 + 4)(x + 1)} dx$$

© MATSEC (University of Malta)

Task 2



A cylindrical can, able to contain half a litre of drink, is to be manufactured from aluminium. The volume of the can must therefore be 500 cm^3 .

- * Find the radius and height of the can which will use the least aluminium, and therefore be the cheapest to manufacture. (i.e., find out how to minimise the surface area of the can).

State clearly any assumptions you make.

- * What shape is your can? Do you know of any cans that are made with this shape? Can you think of any practical reasons why more cans are not this shape?

© Shell Centre for Mathematical Education (University of Nottingham)

Task 3



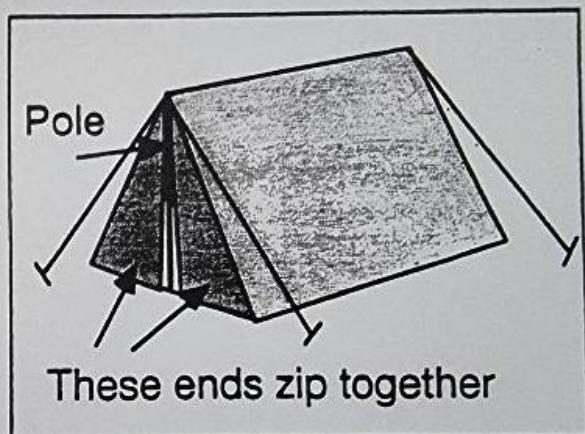
The circular glass top of your neighbor's coffee table breaks. Your neighbor is very upset and would like to replace the glass top but does not know exactly how big it was. He brings you a piece of broken glass that contains part of the boundary of the original top. Describe exactly what you would do in order to figure out the exact size of the original glass tabletop.

© Balanced Assessment Project (University of California, Harvard University, Michigan State University and University of Nottingham)

Task 4

Your task is to design a tent like the one in the picture.

It must be big enough for two adults to sleep in



Show how you will cut the material to make the tent. Show all the measurements clearly.

© Malcolm Swan (University of Nottingham)