

University of London: Institute of Education  
Education and International Development

**The roles of assessment in primary  
education:  
an Egyptian case study**

ELEANORE HARGREAVES

Submitted in fulfilment of the  
Degree of Doctor of Philosophy

**January 1999**



## ACKNOWLEDGEMENTS

I should like to acknowledge Samia Aly who welcomed me into her classroom and her home, during my field work in Egypt; who helped me organise events in my research schedule; who, with the assistance of her daughter, Sara, transcribed and translated the Arabic interviews; who also revealed her true perceptions about my work. I am deeply grateful to Samia, since without her continuous and enthusiastic support, my whole research experience would have been less smooth-running, less insightful and less rewarding.

I am also indebted to El Nasr Girls College, the case study school in Alexandria. The headmistresses, Inaam, Qismet and Samia Salem made me very welcome at school, gave me much information and allowed me free access to classrooms. The teachers, pupils and their parents were also extremely co-operative which ensured that interviews and observations were pleasurable as well as useful.

I am grateful to Jaida, who helped me with my Arabic even though she was working towards her final Certificate of Secondary Education examination at El Nasr Girls College. I am also indebted to Steve Sizmur who spent many hours reading and discussing my work, both at the earliest and the latest stages.

I should like to acknowledge the financial support of the Overseas Development Administration, that provided for me to study full-time for one year. I am also indebted to the National Foundation for Educational Research who allowed me to take the full-time year away from the Foundation, and who supported my second year financially.

Finally, I express enormous gratitude to Pr. Angela Little who has given me great academic guidance in the past three years, has given me the space and opportunities to explore my subject widely, and inspired me to take on ever more exciting challenges. I also express my deep appreciation of Tim for all his personal involvement in bringing this thesis to fruition.

## ABSTRACT

This thesis provides an analysis of the roles fulfilled by educational assessment in primary education. This analysis is called for, now that many policy-makers hold high hopes for the improvement potential of educational assessment.

The field research discussed in this thesis derives from Egypt as a political context and an Egyptian primary school as a case study of practice. From these particular arenas, and from literature relating to assessment, conclusions and implications are drawn up which have relevance to other less-industrialised countries and indeed more-industrialised countries, about the purposes and effects of assessment.

In this thesis, purposes for assessment are defined as either stated or unstated; effects of assessment as either intended or unintended. It is suggested that policy makers' official purposes for assessment may be certification or selection; accountability or system improvement; or the improvement of the individual's learning.

Government purposes for assessment in Egypt have focused more on selection than certification; more on accountability than system improvement; and more on all of these than on the processes of learning in the classroom. Longer-term purposes have also included political or economic improvement.

This research shows that pupils, parents and teachers depend on certification and selection for their personal, social and economic status. As a result, examinations for certification and selection take on prime importance in their daily lives. They may also be associated with pressure and anxiety, a restricted social life and a sense of under-achievement and powerlessness, as well as limiting curriculum and teaching methods at school.

A prime implication from this research is that assessment could improve the processes of learning and teaching, at the same time as fulfilling summative purposes such as certification and selection. The data imply that learning benefits: when pupils are motivated by tests and examinations; when assessment material in tests and examinations is of high quality; finally, when classroom assessment is skilled.

# TABLE OF CONTENTS

	Page
LIST OF TABLES WITHIN THE TEXT	14
THE THESIS. The roles of assessment in primary education: an Egyptian case study	15
<b>PART I. Literature and methods</b>	
<b>CHAPTER 1. Introduction: the roles of assessment</b>	<b>15</b>
<b>1.1 Issues, rationale and emphases</b>	<b>15</b>
1.1.1 The main research question	15
1.1.2 Rationale for the main research question	15
1.1.3 Personal rationale	18
1.1.4 Emphases in this thesis and their rationale	22
<i>Assessment in less-industrialised countries</i>	22
<i>Assessment relating to the primary phase</i>	23
<i>Assessment in practice rather than policy</i>	24
<b>1.2 Research questions</b>	<b>25</b>
<b>1.3 Definitions</b>	<b>26</b>
1.3.1 Towards a working definition of assessment	26
<i>Assessment as information gathering</i>	27
<i>Assessment as testing means or instrument</i>	27
<i>Assessment as measurement</i>	28
<i>Assessment as judgement or evaluation</i>	28
<i>Summary: working definition of assessment</i>	29
1.3.2 Working definition of role	29
1.3.3 Working definition of primary education	29
1.3.4 Working definition of less-industrialised country	30
<b>1.4 Structure of the thesis</b>	<b>30</b>

## **CHAPTER 2. The roles of assessment in primary education: literature review**

<b>2.1 Introduction</b>	<b>32</b>
<b>2.2 The role of assessment for certification and selection</b>	<b>32</b>
2.2.1 Assessment for certification and selection: its contribution to improvement	32
2.2.2 The role of assessment for certification <i>Certification by examination</i>	33 34
<i>Certification through alternatives to examinations</i>	35
2.2.3 The role of assessment for selection	36
<i>The purpose and means of assessment for selection</i>	36
<i>Effects of selection in productivity</i>	38
<i>Effects of selection in social elitism</i>	39
<i>Effects of selection in social mobility</i>	40
<i>The role of objectivity in assessment for selection</i>	41
<i>Some negative effects of examinations</i>	43
<i>Improvement and assessment for certification and selection through improved examinations</i>	45
<b>2.3 The role of assessment for accountability and system improvement</b>	<b>47</b>
2.3.1 Accountability in relation to improvement	47
2.3.2 Monitoring	48
2.3.3 Intentions for system improvement and their realisation	50
2.3.4 The effects on teachers, of monitoring for accountability and system improvement	52
2.3.5 The effects of assessment for accountability and system improvement, on pupils	56
<b>2.4 The role of assessment for the improvement of the individual's learning</b>	<b>58</b>
2.4.1 Assessment for the improvement of the individual's learning as an official, stated purpose	58
2.4.2 Formative and summative purposes	59
2.4.3 Individual learning and assessment which helps improve it	60
<i>The behaviourist and the cognitivist models of learning</i>	62
<i>The transmission model, constructivist model and sociocultural model</i>	64
<i>Surface-level and deep-level learning</i>	68
<i>Basic skills and facts and higher order skills</i>	69
2.4.4 Assessment as motivation for learning	71
<b>2.5 Literature overview: the roles of assessment in primary education</b>	<b>73</b>

<b>CHAPTER 3. Research methods</b>	<b>75</b>
<b>3.1 Introduction</b>	<b>75</b>
<b>3.2 Research questions and methods: rationale</b>	<b>75</b>
3.2.1 Appropriate methods of research	75
3.2.2 Research methods in relation to research questions	77
3.2.3 Main elements of research methods used	80
<b>3.3 Research diary</b>	<b>80</b>
<b>3.4 Case study approach: the case study school</b>	<b>81</b>
3.4.1 The appropriateness of the case study approach	81
<i>Goals of the case study approach</i>	81
<i>Uses of the case study approach</i>	81
3.4.2 The appropriateness of the case study school	82
<b>3.5 Pilot study</b>	<b>84</b>
<b>3.6 Analysis of writing</b>	<b>85</b>
3.6.1 Aims for writing tasks	85
3.6.2 Sentences to complete	86
3.6.3 Time schedules to fill in	87
3.6.4 Examination essays	87
<b>3.7 Interviews</b>	<b>88</b>
3.7.1 Overview of the interview samples	88
3.7.2 Aims for the interviews	88
3.7.3 Pupil interviews	88
<i>The pupil interview process</i>	88
<i>The pupil interview sample</i>	89
3.7.4 Parent interviews	90
<i>The parent interview process</i>	90
<i>The parent interview sample</i>	91
3.7.5 Teacher interviews	91
<i>The teacher interview process</i>	91
<i>The teacher interview sample</i>	92
3.7.6 Interviews with headteachers	93
3.7.7 Interview with the inspector	94
3.7.8 Interviews at NCEEE	95
3.7.9 Interview at the Ministry of Education	96
3.7.10 Interview with USAid curriculum development trainer	97
<b>3.8. Observation</b>	<b>97</b>
3.8.1 Classroom observations	97
3.8.2 Observations of examination invigilation	99
3.8.3 Observations of the marking of examinations	99
<b>3.9 Validation</b>	<b>99</b>
<b>3.10 Issues concerning the research practice</b>	<b>100</b>

## **PART II. The research findings**

<b>CHAPTER 4. The roles of assessment in Egypt from 1798 to 1996</b>	<b>102</b>
<b>4.1 Introduction</b>	<b>102</b>
<b>4.2 The historical role of assessment in Egypt</b>	<b>103</b>
4.2.1 The government's earliest aims for assessment	103
4.2.2 The instrumental reasons for the first primary schools	104
4.2.3 Revolutionary aims for assessment	105
4.2.4 Tensions between Nasr's political ideology and the economy	106
4.2.5 Sadat's shift in aims for assessment	107
4.2.6 Mubarak's aims	108
<b>4.3 The current system in Egypt</b>	<b>110</b>
4.3.1 Primary education provision	110
4.3.2 Teaching at primary level	111
4.3.3 Selective examinations	111
4.3.4 Non selective examinations	114
<b>4.4 Improvement of the individual's learning</b>	<b>114</b>
4.4.1 Curriculum reform to improve the learning of the individual	114
4.4.2 Assessment to improve the learning of the individual	115
<b>4.5 Assessment for accountability and system improvement</b>	<b>117</b>
<b>4.6 Overview: the roles of assessment in Egypt from 1798 to 1996</b>	<b>119</b>

<b>CHAPTER 5. The roles of assessment in the case study school</b>	<b>121</b>
<b>5.1 Introduction</b>	<b>121</b>
<b>5.2 The case study school</b>	<b>121</b>
5.2.1 The case study school in the context of national school types	121
5.2.2 Structure of the case study school	123
<b>5.3 The role of assessment for certification and selection in the case study school</b>	<b>124</b>
5.3.1 Certification as curriculum coverage	124
5.3.2 The means of assessment for certification and selection	125
5.3.3 The purpose of assessment for certification and selection	128
<i>The selective function of assessment</i>	128
<i>Competition to achieve high marks</i>	129
<i>The role of objectivity in assessment for selection</i>	130
<i>Selection and social elitism in the case study school</i>	132
<i>The negative role of assessment for selection in learning and teaching</i>	135
<b>5.4 The role of assessment for accountability and system improvement in the case study school</b>	<b>138</b>
5.4.1 Assessment for accountability or improvement in the case study school	138
5.4.2 The means of assessment for accountability	140
<i>Assessment for accountability through examination results</i>	140
<i>Inspection by Ministry inspectors</i>	143
<i>Monitoring for accountability through internal inspection</i>	145
<b>5.5 The role of assessment for the improvement of the individual's learning in the case study school</b>	<b>147</b>
5.5.1 Perspectives on learning in the case study school	147
5.5.2 Specific assessment tasks given in class for formative purposes	149
<i>Oral quizzing</i>	149
<i>Blackboard questions</i>	149
<i>Whole class recitation</i>	150
<i>Classwork and marking or supervising</i>	150
<i>Homework and marking</i>	150
<i>Weekly tests</i>	151
<i>Monthly tests</i>	151
5.5.3 Assessment as motivation for learning	152
<b>5.6 Overview: the roles of assessment in the case study school</b>	<b>153</b>



**CHAPTER 6. The roles of assessment for pupils 155**

**6.1 Introduction 155**

**Section I. Overview: pupils' views and feelings, as expressed in writing**

**6.2 Overview of pupils' views and feelings 155**

6.2.1 Pupils' positive feelings about their schooling 155

6.2.2 Pupils' ambitions for after schooling 156

6.2.3 The purpose and use of examination results 157

6.2.4 Pupils' negative feelings during the examinations 158

6.2.5 Pupils' views on how to succeed 159

**6.3 The 24 hour diary record 159**

**Section II. Detail: pupils' views and feelings, as expressed in interview**

**6.4 Interview pupils' views and feelings about schooling 161**

6.4.1 Pupils' feelings about their school 161

6.4.2 Pupils' ambitions for after schooling 162

**6.5 Interview pupils' views on the nature of examinations 163**

6.5.1 The purpose and use of examinations 163

6.5.2 The origins of examinations 163

6.5.3 The easiest and hardest examination subjects 165

**6.6 Interview pupils' experiences of examinations 167**

6.6.1 Preparing for examinations 167

*Hours of study during the week* 167

*Help at home* 168

*Study at weekends* 169

*Private lessons* 170

6.6.2 Feelings during the examinations 171

6.6.3 Rewards and punishments 173

**6.7 Overview: the roles of assessment for pupils 175**

<b>CHAPTER 7. The roles of assessment for parents</b>	<b>176</b>
<b>7.1 Introduction</b>	<b>176</b>
<b>7.2 Parents' perspectives on the EGC</b>	<b>176</b>
7.2.1 The educational background and employment of parents	176
7.2.2 The importance of the school's good academic standards	177
7.2.3 The importance of non academic aspects of EGC	178
7.2.4 The importance of the English medium in school	178
<b>7.3 Parents' support to children taking examinations</b>	<b>180</b>
7.3.1 Supporting children in preparation for examinations	180
7.3.2 Supporting children during the examinations	183
<b>7.4 Parents' perspectives on examination results</b>	<b>184</b>
7.4.1 The value of good examination results to parents	184
7.4.2 Parents' reactions to pupils' examinations results	185
7.4.3 Parents' views on the validity of examination results	186
7.4.4 Parents' hopes for their daughters' future	<b>187</b>
<b>7.5 Parents' perspectives on the primary assessment system</b>	<b>188</b>
7.5.1 Parents' views on how the assessment system could be improved	188
7.5.2 Parents' awareness of recent reforms to the assessment system	190
7.5.3 How the system had changed since parents' own school days	192
<b>7.6 Overview: the roles of assessment for parents</b>	<b>194</b>

<b>CHAPTER 8. The roles of assessment for teachers</b>	<b>197</b>
<b>8.1 Introduction</b>	<b>197</b>
<b>8.2 Teachers' place in the system</b>	<b>197</b>
8.2.1 Teachers' training and reasons for teaching	197
8.2.2 Teachers' general feelings about teaching	199
8.2.3 Teachers' views about EGC's status and teaching in the English medium	200
<b>8.3 Teachers' views on examinations</b>	<b>201</b>
8.3.1 Teachers' views on the purpose of primary education and examinations	201
8.3.2 Teachers' views on the necessity of examinations	203
8.3.3 Teachers' views on why examination results were important	204
8.3.4 Teachers' perceptions about private lessons	206
8.3.5 Teachers' views on examination validity	208
<b>8.4 Teachers' views on primary assessment reforms</b>	<b>209</b>
8.4.1 Teachers' awareness of central reforms and views on the reforms	209
8.4.2 Teachers' views on how the primary assessment system could be improved	212
<b>8.5 Overview: the roles of assessment for teachers</b>	<b>213</b>
8.5.1 Teachers' reasons for wanting to maintain the system	213
8.5.2 Maintaining the system by default	215

## PART III. Conclusions and implications

### CHAPTER 9. The roles of assessment in primary education: some conclusions and implications 216

#### 9.1 Introduction 216

#### Section I. The roles of assessment: conclusions from the research

#### 9.2 The roles of assessment for different groups 218

9.2.1 The roles of assessment for pupils, parents and teachers respectively 218

9.2.2 The roles of assessment in policy and practice respectively 220

#### 9.3 Assessment for certification and selection: selection of an elite to employment and competitive pressure and anxiety for individuals 221

9.3.1 Selection for employment 222

9.3.2 Selection of an elite 222

9.3.3 Pressure to succeed and its effects 223

#### 9.4 Assessment for certification and selection: normalisation through examination objectivity and underachievement for most 225

9.4.1 Objectivity 225

9.4.2 The normalising function of assessment 226

#### 9.5 Assessment for certification and selection: social time spend studying 228

#### 9.6 Assessment for accountability and system improvement: unity or conformity and a sense of powerlessness for individuals 230

9.6.1 Unity 230

9.6.2 Conformity 230

#### 9.7 Assessment for the improvement of the individual's learning: official purpose and actual practice 232

9.7.1 Assessment for the improvement of the individual's  
learning as an official assessment purpose 232

9.7.2 The need for assessment for the improvement of  
the individual's learning 233

#### Section II. The roles of assessment: implications from the research

#### 9.8 Implications for assessment: assessment for any purpose must improve learning and teaching 235

9.8.1 Assessment for the improvement of the individual's  
learning: implications for English practice 239

#### 9.9 Implications for research 240

9.9.1 Carrying out research in a less-industrialised country 241

9.9.2 Further research into the roles of assessment 241

9.9.3 Further research into learning theory 242

9.9.4 Further research with primary school pupils 242

#### 9.10 This thesis: assessment and learning 244

## **PART IV. Appendices and references**

<b>Appendix A: schedules</b>	<b>245</b>
Schedule 1. Pupil interview schedule	245
Schedule 2. Parent interview schedule	245
Schedule 3. Teacher interview schedule	246
Schedule 4. Classroom observation schedule	246
<b>Appendix B: tables</b>	<b>247</b>
Table 1 to show pupil interview sample	247
Table 2 to show personal qualities of pupils in interview sample	248
Table 3 to show parent interview sample	249
Table 4 to show teacher interview sample	250
Table 5 to show parents' educational background and current employment	251
Table 6 to show number and percentage of both primary school type in Egypt, 1995-6	252
Table 7 to show numbers of classes and pupils at each stage in El Nasr Girls College (EGC) in 1995-6	253
Table 8 to show curriculum distribution: grades 2, 3 and 4, by number of lessons	254
<b>References</b>	<b>256</b>

## LIST OF TABLES WITHIN THE TEXT

<b>Table</b>	<b>Information</b>	<b>Page</b>
3.1	Observations of classes	98
4.1	Numbers of schools and pupils at different school phases in Egypt, 1995-6	110
5.1	Structure of the administration of written assessments	127
6.1	Overview: Pupils' favourite aspects about their schooling	156
6.2	Overview: Pupils' ambitions for after schooling	157
6.3	Overview: Pupils' views on the purpose and use of examination results	158
6.4	Overview: Pupils' negative feelings during the examinations	159
6.5	Overview: Pupils' views on how to succeed	159
6.6	Summary of grade 4 pupils' daily schedule for 24 hours	160
6.7	Interview pupils' feelings about school	162
6.8	Interview pupils' ambitions for the future	162
6.9	Interview pupils' views on the purpose and use of examinations	163
6.10	Interview pupils' beliefs about where examinations came from	165
6.11	Interview pupils' on easiest and hardest examination subjects	166
6.12	Interview pupils' average hours of study on a week night	167
6.13	Interview pupils' family members who usually helped them study	168
6.14	Interview pupils' weekend leisure activities	169
6.15	Interview pupils' feelings before and during examinations	172
6.16	Rewards interview pupils received	173
6.17	Punishments interview pupils received	174
7.1	Parents' reasons for choosing EGC for their daughters	179
7.2	Private lessons described by parents	182
7.3	How parents said they prepared children for examinations	183
7.4	Reasons why examination results were important to parents	185
7.5	Parents' suggestions for reforms to the primary school system	190
7.6	Central reforms described by parents	192
8.1	Teachers' reasons for becoming a teacher	199
8.2	Teachers' reasons for enjoying teaching	200
8.3	Teachers' views about the status of EGC and the English medium	201
8.4	Teachers' beliefs about the purpose of primary education	202
8.5	Teachers' views on the necessity of examinations	204
8.6	Teachers' views on why examination results were important	206
8.7	Teachers' perceptions about private lessons	207
8.8	Teachers' views on examination validity	209
8.9	Teachers' comments on central reforms	211
8.10	Teachers' views on how the primary assessment system could be improved	213

# **PART I. Literature and methods**

## **Chapter 1. Introduction: the roles of assessment**

### **1.1 Issues, rationale and emphases**

#### **1.1.1 The main research question**

The main research question addressed by this thesis is, what are the roles of educational assessment in primary education? The issues this thesis addresses are: What roles does assessment fulfil for different groups of individuals? How is educational assessment used politically and socially? How is it experienced socially and by those within formal education? In particular, how do policy-makers intend assessment to be used by individuals and how do individuals experience and use assessment in actual practice?

The thesis distinguishes three official purposes for assessment: for the certification and selection of individuals; for accountability and system improvement; for the improvement of the individual's learning. It then explores how assessment for these stated purposes also fulfils purposes that may or may not be stated and affects the lives of individuals in ways that may not be intended.

The main research question is which roles educational assessment fulfils in primary education. In this thesis, however, the question is addressed in the context of the Egyptian assessment system. It describes the assessment system in Egypt and describes an Egyptian primary school as a case study. These descriptions then provide for an analysis of the roles of assessment in primary education more generally.

#### **1.1.2 Rationale for the main research question**

Frederikson and Collins (1989) claimed that, 'The goal of assessment has to be, above all, to support the improvement of learning and teaching' (p.32). Assessment in education has recently gained international prominence and its role is being reconsidered by individuals, schools and, especially, states across different areas of the world, in particular because of its perceived potential for improvement: in the immediate term, for the improvement of educational systems and individuals' learning; but in a broader sense, for social and economic improvement, too. While '... it is hardly a novelty for testing and assessment to figure prominently in policy-makers' efforts to

reform education' (Linn, 1993, p.1), in the current *renaissance*, there has been a shift in emphasis in relation to the roles of assessment. The titles of four books which were particularly useful in informing this thesis are indicative of the international recognition of new roles for assessment in the 1990s: Berlak's (1992) book, *Toward a new science of educational testing and assessment*; Broadfoot, Murphy and Torrance's (1990) book, *Changing educational assessment*; Little and Wolf's (1996) book, *Assessment in transition*; Gipps' (1994) book, *Beyond testing*.

A recognition of assessment's potential roles is increasing in many places, and its focus is shifting everywhere (Dwyer, 1990). Both increase and shift in the emphasis on assessment's role were illustrated by Lockheed (1996) who reported World Bank support for assessment world-wide: no assessment projects were funded before 1975, while by 1990, 40 per cent of educational funding was allocated to assessment projects. Within that time, there was an increase in projects supporting assessment for certification and selection; and a far greater increase in projects for monitoring as well as for teacher diagnosis of student learning. Of particular relevance to this thesis is the fact that in all cases, most support went to the basic stage of education where assessment strategies dedicated to the younger children were now being explored.

These figures do not simply reflect World Bank policy. The heightened profile of assessment and its shifting role was stressed by Berlak (1992) writing about the United States, and Murphy and Torrance (1988) writing about England: each wrote of the emergence of assessment out of the classroom into the political limelight during the 1980s. Little and Wolf (1996) described the shift of emphasis as 'A move from access and quantity of education to outcomes and quality' (p.vii); Gipps (1994) described it as a '... paradigm shift, from psychometrics to a broader model of educational assessment' (p.1). Speaking specifically about less-industrialised countries, Kellaghan and Greaney (1992) described '... a broadening of the scope of assessment and ... a shift in responsibility for assessment' (p.65).

Because of this increase and shift in emphasis on assessment, the issue of its role, that is, how assessment may be used by policy-makers, or used and experienced by those who are subject to policy, has become more relevant as well as more complex. Suddenly educational assessment is seen as a powerful means to achieve a variety of



educational or other goals, it is seen to fulfil a variety of roles: but whose goals and which roles? Are these the goals of policy-makers, pupils, parents or teachers? How do we know that stated or 'official' roles for assessment are the roles it really fulfils? Or does it fulfil its stated roles and other roles as well, which suit the goals of some parties involved but are destructive to others' (Little, 1990; Wolf, 1996)?

If assessment is to be used to achieve specific goals, then all the roles it fulfils in achieving those goals could usefully be uncovered and made explicit: that is, what actually happens as a result of a specific policy when it is implemented in practical as well as social terms; how different parties working within a particular assessment policy use and experience assessment; whether they are using assessment for their own explicit aims or being more subtly affected by the assessment; whether immediately or in the longer term. In particular, assessment's role in 'improvement' could be examined in terms of what is actually being improved.

The complex relationship between stated policy and how various parties use and experience that policy, is exemplified by Little and Wolf (1996):

Governments, in particular, want to monitor, and also to control education, in order to ensure 'value for money'. At first sight, this objective, and that of promoting learning, may seem closely linked, since both are concerned with individuals' achievement. In practice,... tensions between the two quickly become evident (p.vii).

Lee and Ninnes (1995) explained the complex relationship between official purposes and actual effects as lying '... in different assumptions held about the purposes of education by educational theorists and practitioners on the one hand, and clients of the system itself on the other' (p.174). Dwyer (1990) suggested that one reason for its complexity was the subsidiary role played by educational experts when policy was made: that is, policy-makers may not have enough 'inside' information to predict potential impacts. More information about the variety of effects a specific assessment policy could have, would therefore be useful. In the words of Eckstein and Noah (1993), '... policy-makers constantly need to be mindful of the likelihood of unintended, undesirable consequences of their decision' (p.16).

On the other hand, policy-makers may be fully aware of potential effects yet express official policy in such a way to allow the implementation of unpopular policies: that is,

they may disguise the true intended use of the assessment. In the words of Black (1993), 'A new policy statement has to be one that can be 'sold' to the public' (p.425). Torrance (1996) suggested that the more explicit the intentions of a policy, the more likely it is to be used in the way policy-makers intended. This is because assessments depend for effect on how recipients perceive them. But it may or may not be policy-makers who convey purposes and so determine how people respond to them. Corbett and Wilson (1988) claimed:

... if students, teachers, or administrators believe that the results of an examination are important, it matters very little whether this is really true or false - the effect is produced by what individuals perceive to be the case (p.28).

More information about how recipients may perceive assessment policies, or the role they fulfil for various parties despite official purposes, would therefore be useful if assessment policy is to be implemented effectively.

By researching the roles of assessment, a clearer understanding can be reached of assessment's officially stated use, its actual use and its obvious effects as well as subtler ones. This thesis is, therefore, jumping on the assessment bandwagon, celebrating its prominence, tapping into the optimism about its potential and addressing the need for research: so that appropriate decisions may be made which take full account of assessment's potential roles in different contexts, based on extensive and in depth data. This thesis aims to make a contribution towards such an extension and deepening of understanding about the roles of assessment.

### **1.1.3 Personal rationale**

In this section, an account is given of how assessment impinged on my own life and how this thesis consequently began. Assessment in the form of written examinations was a dominant part of my own school experiences. On leaving school in 1979, I went to work for four months in the Egyptian primary school which is the focus of this research. There, I invigilated examinations for seven year olds which were as formal as the A Levels I had just taken.

This first experience in Egypt inspired my subsequent interest in less-industrialised countries. After university, I trained to be a primary teacher so that I could return to Egypt and teach professionally. I trained in the days before a national curriculum and

national assessment had been introduced in England. In 1984, I was able to return to teach in the Egyptian primary school, teaching to the Egyptian examinations prescribed by the Ministry. I later became an infant teacher in a small, multigrade British school in Alexandria where I devised my own assessments. These early teaching experiences confirmed my concern for the primary phase as fundamentally important in children's educational and social progress.

I took two years out in 1988 to complete my Masters in Education degree in the Republic of Ireland. On returning to the school scene in England, I found the national curriculum and its assessment being imposed at infant level, or Key Stage 1. My work then was with an educational charity. With this charity, I developed curriculum materials which helped teachers in London teach each Statement of Attainment at Key Stage 1 in the new curriculum, in preparation for the Standard Assessment Tasks of summer, 1991.

Within two years, I had moved jobs and was myself involved in writing and researching those Standard Assessment Tasks, now working with the National Foundation for Educational Research in England and Wales. We developed the early assessments of English, mathematics and science for seven year olds (Sainsbury, 1996).

Like many teachers, I had not thought very much about assessment or why it was important. It had not been formalised at infant level in this country before. Now my entire job was dedicated to developing national tasks and test items, trialling them with groups of teachers and with a national sample, discussing them with officers at the government assessment agencies, first at the School Examination and Assessment Council then at the School Curriculum and Assessment Authority, writing instructions so that teachers could implement the assessments reliably and, finally, evaluating the national process and its outcomes. I had to accept that assessment now had a prominent official stated role in the English education system.

In 1994, I spent two months in Delhi, India, attempting to devise an assessment system for two large and impoverished pre-schools. I had to apply my insights about assessment to a very different context and carry out my research in some challenging situations. This experience alerted me abruptly to the role of context in the practice of assessment. On returning to England, I found I was more objective about national

curriculum assessment and the more formal role it had by now. In particular, national evaluations that I worked on, including interviews with teachers and children (Sainsbury *et al*, 1994; Ashby *et al*, 1995), observations of the trialling of assessment materials, and my own analysis of children's test scripts (Hargreaves, 1995) made me question the effects that nationally imposed assessment was having in practice. These exercises made me see that its official role was not necessarily the role it fulfilled for other parties involved.

One year 2 teacher I interviewed during a national evaluation brought this lesson home to me. I asked the teacher several questions about the sorts of changes he would like to see to arrangements for administering the national tests. For example, how did this test compare with last year's? Were the administration instructions clear? Was the test too long? After a while, he told me to stop and explained that I was missing the point. He told me the questions should be about what kind of *assessment* he considered appropriate, not what kind of arrangements: and nationally imposed tests would not be a kind of assessment he would consider appropriate at all. He felt that formative, classroom assessments were more beneficial for improving pupil learning.

Interactions like these led me to question further the official purposes for the assessments. I saw these as primarily monitoring for accountability as well as system improvement in the immediate term and national economic improvement in a broader sense: I saw a mismatch between these purposes and how the assessments were actually used and experienced by children and teachers whose prime concern was to enhance each child's own learning (Hargreaves, 1996). I asked myself whom this accountability was benefiting, what was actually being improved within the system and who would gain in the longer term. At this point I decided to get a better grasp of the potential roles of assessment and explore the different roles it had for different groups of people.

I turned my thoughts to Egypt whose primary system I felt had always been very unlike the *laissez faire* English system and I asked what roles assessment had for different parties there. I had the following image vividly in my head, a picture of the grade 3 children I had taught in Egypt:

It is their final examination in examination week in the hot month of May. They are sitting in examination halls at separated rows of wooden desks, divided among children from other classes whom they do not know. Each has a special examination number. The teachers invigilating are restless but severe, clamping down on any child who fidgets. But all the children have finished writing and the papers have been collected in and signed by the senior teacher.

Suddenly the finishing bell rings. Without waiting to be dismissed, over three hundred grade 3 children pour wildly out of the buildings, some through the windows, others cramming through the doors. They tear screaming across the playground with their heavy school bags flying behind them. Across the playground, parents are waiting to pick them up, eager to know how their child performed in the final examination, but also celebrating the end. The children put on casual clothes or party dresses and are driven away for four months of study-free summer.

At the time I witnessed this scene, I had just stood and enjoyed the children's ebullience. But now I questioned. How far did these examinations benefit the children? Did children not suffer because of them in Egyptian society? Or did Egyptian parents consider them socially beneficial? What benefits did teachers see in them, did they perhaps use them to boost their own status? Was the current role assessment fulfilled subtly beneficial to some people and subtly destructive to others? How exactly did individuals experience assessment there in practice? And which official roles was this assessment system fulfilling in these children's lives? Were they the traditional roles of selection and employment, and if so, what was their relevance at primary level? Did they share anything in common with assessments elsewhere in the world? Did they share any of the intentions behind the new assessments in England? What other kinds of assessment and purposes for assessment could have been more beneficial? So my thesis began from these initial questions.

#### **1.1.4 Emphases in this thesis and their rationale**

The personal details given above should help explain the focus of this thesis and also the particular emphases within it. In this thesis, the role of assessment is explored with especial emphasis on how assessment is used and experienced in less-industrialised countries of the world; how it is used and experienced at primary level; also, how it is used and experienced in practice as well as in policy.

##### ***Assessment in less-industrialised countries***

Little (1996a) observed that *international* research data tended actually to be dominated by research of the industrialised northern countries:

Researchers whose primary national identities lie within countries of the North undertake research in the North and the South. Researchers whose primary national identities lie within countries of the South undertake research mainly in the South (p.5).

This thesis falls into the category of research undertaken by a northerner in the south. As Little observed, research from a southern point of view undertaken in a northern context is rare and this thesis does not lessen this imbalance. But the thesis does aim to contribute towards lessening the northern domination of research documentation about assessment. By analysing in depth how assessment is used and experienced in Egypt, it provides a new insight into how assessment is used and experienced in a less-industrialised country and this insight may help describe and have implications for assessment in other less-industrialised countries or indeed more-industrialised ones.

The thesis aims to be of particular value in that, unlike much available assessment literature about the less-industrialised world (for example, De Luca, 1994; Loxley, 1983), it emphasises how individuals use and experience assessment in their everyday lives, as well as how policy is implemented. In this way, light is thrown on how assessment policy in a less-industrialised country is actually experienced in practice, not only in theory (but see Brooke and Oxenham, 1980; Little, 1985; Payne and Barker, 1986). It is valuable to give a northern audience a particularly vivid picture of a southern situation.

The thesis also describes Egyptian assessment policy itself and this is little documented in the English language (but see Carroll, 1995; Cochran, 1986; Ministry of Education

[MoE], Egypt, 1992 and 1995). Since educational research, especially qualitative research, is very limited in Egypt, there is little documented about assessment policy or practice even in Arabic. So while the thesis essentially explores the role of assessment *per se*, it does not only feed into an international debate about the roles of assessment; outcomes relating to the Egyptian context contribute to a limited body of literature about assessment in Egypt; the emphasis on individuals provides for more research evidence to reach this limited area, too. Chapter 3 (Research Methods) provides details about the particular circumstances of carrying out research in Egypt as a less-industrialised country (see also Vulliamy *et al*, 1990).

### ***Assessment relating to the primary phase***

Assessment research relating to primary education has been considerably scarcer than research relating to secondary or tertiary assessment, even though the primary phase of education could be considered fundamental to the entirety of each child's future experiences. Despite the UNESCO-UNICEF emphasis on primary education as 'the cutting edge' of any Education For All strategy (Chinapah, 1996, p.45), and despite World Bank assessment funding for basic education exceeding funding for any other phase (Lockheed, 1996), research into assessment often focuses on secondary or tertiary education because of assessment's traditionally dominant role in selecting for employment (Dore, 1976). Hallak (1990) wrote:

Restoring economic growth, which inexorably elbows its way to top priority, still means for many policy-makers ... concentrating the effort on secondary and tertiary level ... with scant attention being paid to the growth or improvement of primary schooling (pp.89-90).

Forms of assessment suitable for secondary or tertiary education may be 'uncritically extended downward' (Madaus, 1991) and primary teachers may see the primary years as preparation for selection assessment in the senior years (Bachor and Anderson, 1994). In these senses, as the role of assessment is extended away from just selection, its role in primary school becomes an issue which needs to be addressed independently of higher phases of education.

Research on assessment at primary level which is based on interview data has also been inhibited because, perhaps owing to certain theories of child development, there has been a limited view of children's ability to reflect on and talk about the adult world

(Pollard *et al*, 1997). By choosing interview and observation as research methods for this thesis, it was shown possible to record the responses of children in a primary school and so contribute to the limited literature based on children talking, whether in English, Arabic or another language. In this way, primary school children could add their own insights about primary schooling and assessment, and these insights complemented parents' and teachers' information in a unique way.

In summary, this thesis extends assessment literature about the primary phase by clarifying how assessment is used and experienced by individuals in a primary school so that implications may be drawn out as to specific features of its role in the primary phase.

### ***Assessment in practice rather than policy***

The new emphasis on and extension of the role of assessment referred to above, often relates to assessment's role in educational policy: it relates to planning bodies' *intentions* for outcomes such as improvement, and the processes they put in place to achieve these, rather than to the *effects* of assessment and the experiences which result from it. Vulliamy *et al* (1990) suggested,

... there has been a tendency for the literature on education in developing countries to be concerned more with a discussion of policies and system-wide features than with observation and the realities of schooling at the chalk face... (p.17).

Recent research which does attend to the effects of assessment tends to focus more on educational than social effects (Alderson and Wall, 1993; Bracey, 1983; Little, 1984; Madaus, 1991; Somerset, 1996) and more on its effects on societies than on individuals socially (Berlak, 1992; Broadfoot, 1996a; Dore, 1976; Heyneman and Ransom, 1990; Lolwana, 1996; Lulat, 1988; Meadmore, 1993).

The research described and analysed in this thesis therefore focuses expressly on individual pupils, parents and teachers (see also Ames, 1992; Little, 1978; Payne and Barker, 1986; Seginer, 1983), on how assessment is actually practised in their school and how they experience and perceive assessment in a social context. The thesis aims to '... uncover in the everyday operation of testing a series of well-concealed and mostly unintended consequences that exercise far deeper and more pervasive influence in social life than is commonly recognized' (Hanson, 1993, p.2 quoted by Madaus,



1994, p.222). Parents were included in the study because of their more distant connection with assessment and because its effects on them may be less recognised. The aims of focusing on individuals, through observation and interview, are: to gain a clear picture of what really happens in practice; to explore the social effects this has on individuals; and to deepen understanding of the social meaning assessment has for individuals. It was these aims which helped form the four subsidiary research questions as described below.

## **1.2 Research questions**

The research question addressed by this thesis is, what are the roles of educational assessment in primary education? The original research for this thesis was carried out in a primary school, El Nasr Girls College Junior School in Alexandria, Egypt. This school provided for the case study. The title of the thesis is therefore, 'The roles of assessment in primary education: an Egyptian case study'.

The four subsidiary research questions which underpinned this field research were as follows. The answers to these questions contributed to the more general analysis of the roles of educational assessment in primary education.

1. What were assessment's official roles in Egypt generally and in the case study school specifically?
2. In what senses did pupils, parents and teachers experience the assessment system positively?
3. In what senses did pupils, parents and teachers experience the assessment system negatively?
4. What were pupils', parents' and teachers' experiences and perceptions of central reforms to assessment?

By answering the first question, information was to be gathered about how policy-makers intended assessment to be used, based on details about the national and school-based assessment mechanisms. This information was also intended to supplement the limited details available about assessment in Egypt generally. In addition, it provided

the context from which to explore the effects of official policy in the social lives of individual children, parents and teachers and on pupils' and teachers' experiences at school. The last of the four questions was designed to allow a direct juxtaposition of official policy and individuals' experience of it in actual practice. *Reforms* were chosen as the focus for this question as a further exploration of assessment's recently emphasised role for the improvement of the educational system and the individual's learning (Torrance, 1996).

The second and third research questions were designed to produce qualitative data about how individuals used assessment to fulfil their own practical and social purposes and what experiences they went through to achieve these and to accommodate the system. The similarity or difference of view among the three groups of individuals could then be explored, too. The clear-cut categories of *positive* and *negative* were used to allow some open-endedness of response. They were also useful in addressing the issue that national policy, which was intended to have positive effects nationally, need not have positive effects in the lives of individuals within the system.

## **1.3 Definitions**

### **1.3.1 Towards a working definition of assessment**

Cohen (1990) pointed to the ironic situation that, although

... assessment is a central concern of both the educational and political agendas at present... there may exist widely differing perceptions of the *meaning* of assessment... (p.35).

Cohen illustrated the difficulty of arriving at one satisfactory definition of a term, by describing how 600 respondents to an Australian national survey about curriculum defined 'curriculum' in 600 different ways. He concluded that '... there are generally many alternative and equally acceptable definitions of educational terms' (p.34). In the ensuing section, some alternative but equally acceptable definitions of 'assessment' are described. Their variety reflects the variously perceived roles of assessment, which are the subject of this thesis.

### ***Assessment as information gathering***

Maeroff (1991) saw 'assessment' as 'gathering reliable information' in that the root of the word 'assessment' meant

to "sit with" a learner and seek to be sure that a student's responses really mean what they seem to mean (p.276).

Nitko (1995) also defined 'assessment' as 'information gathering':

The term assessment, then, refers only to the process of gathering relevant information, for the expressed purpose of making educational decisions, not to the instrument for gathering it (p.322).

Nitko's definition is particularly useful, if narrow, because in defining assessment as 'information gathering', Nitko distinguished 'assessment' from 'testing', which referred to more or less formal instruments of assessment; from 'measurement', which was the process of reporting a student's performance on a numerical scale; and from 'evaluation', which was the process of judging the quality of a student's performance.

Jones (1995), however, combined three of the four functions differentiated by Nitko into one definition of assessment. Collecting information was only the first of three components. Once collected, assessment information had to be used by teachers to make professional judgements, which is 'evaluation' in Nitko's terms; these judgements related to pupils' progress which was 'measured'; and were then fed back into curriculum and instruction planning, based on further 'evaluation'. Jones' definition is noticeable for omitting to state the 'testing' aspect of assessment, perhaps because she was writing for a testing dominated audience.

### ***Assessment as testing means or instrument***

As Nitko noted, the term 'assessment' is sometimes equated with assessment means, method or instrument. In some parts of the world 'assessment' is equated with 'examinations', but this is because examinations are a dominant form of assessment, not because assessment is an alternative name for examinations. Examinations may therefore be seen as a means or an instrument of assessment. Gipps (1994) defined 'assessment' as 'method', but in broader terms than 'examinations' or 'tests', thus agreeing with Nitko's distinction between formal and informal testing. 'Assessment' meant

... a wide range of methods for evaluating pupil performance and attainment including formal testing and examinations, practical and oral assessment, classroom based assessment carried out by teachers and portfolios (Gipps, 1994, p.vii).

### ***Assessment as measurement***

Futcher (1989) considered the difference between 'assessment' and 'measurement' as a 'dichotomy':

... a major problem: measurement or assessment? So far, the terms have been used interchangeably, but it is clear that they are not in fact interchangeable. Measurement entail [sic] that the method employed be reliable and have both content and construct validity. Assessment, on the other hand, lacks at least one of these three features. At some stage the subjective opinion of the assessor comes into play (p.261).

Gipps (1994) described the distinction between 'measurement' and 'assessment' in terms of a move away from 'psychometric' measurement with its 'highly standardized procedures' which might be equated with Nitko's 'numerical scale' of measurement, towards 'a broader model of educational assessment' (pp.1-2).

### ***Assessment as judgement or evaluation***

Nitko considered 'evaluation' to be the process of judging the quality of a student's performance and to be distinct from 'assessment'. By some definitions, 'assessment' means 'judgement'. For example, Little and Wolf (1996) defined 'assessment' as '... all judgements of educational performance' (p.viii), but their definition extended beyond judgement of the individual student, which was what Nitko implied, to judgements of educational performance

... which are used in individual or aggregated form, for one or more purposes and by a range of persons and institutions (p.viii).

Torrance (1986) distinguished 'assessment' from 'evaluation' by claiming that the former enabled statements about individuals, while the latter enabled statements about the educational service that individuals were receiving. He was dealing with the roles of assessment rather than defining them, but his distinction between the two accorded with Nitko's. It was different, though, in that Torrance referred to evaluation as relating to educational bodies rather than individuals.

### ***Summary: working definition of assessment***

For the purposes of this thesis, the broadest definition of assessment is used. With reference to the descriptions above, and including all possible understandings of the word, the following are selected to constitute a broad working definition. There is a large extent of overlap across the four sub components of the definition and each also stands alone.

- Assessment refers to the **gathering of information** about individuals' educational achievement, or the achievement of larger educational structures such as schools or state bodies
- Assessment refers to the **measurement** of students or educational facilities on a statistical scale, allowing comparisons and standardisations
- Assessment refers to **evaluation or judgements** about individuals' educational achievement or the achievement of larger educational structures such as schools or state bodies
- Assessment refers to **means or instruments of testing**, whether, for example, formal examination or informal discussion, whether for individuals or for larger educational structures.

### **1.3.2 Working definition of role**

Fowler and Fowler (1978) defined 'role' as 'What a person or thing is *appointed* or *expected* to do'. In the context of the role of assessment, this could be rephrased as 'What assessment is intended to do or what happens when it is effected and experienced'.

### **1.3.3 Working definition of primary education**

'Primary education' refers in this thesis to children in the initial years of schooling. In many systems these could be categorised as children in grades 1 to 6 or 7, as infant and junior aged children or as elementary school children. In some countries, primary and basic education denote the same initial years of schooling but in most, basic education includes two or three more years (Meng Hong-wei, 1996). In the Egyptian system, the

span of primary education was grades 1 to 5 and of middle school was grades 6 to 8. All the grades from 1 to 8 were described as 'basic education'.

#### **1.3.4 Working definition of less-industrialised country**

An industrialised country may be said to have three main characteristics: the majority of its labour force are producers of a secondary or tertiary nature; it is constantly attempting to expand productivity; consequently, there is a rapid rate of technological innovation (Aron, 1980, quoted in Broadfoot, 1996a, p.10). In a less-industrialised country, therefore, these features are less likely to apply. That is, in a less-industrialised country, a higher proportion of the labour force are primary producers; it is more traditional and relatively stable in terms of productivity; there is a less rapid rate of technological innovation.

### **1.4 Structure of the thesis**

Part I of this thesis deal with literature and methods. Part II presents the findings from the research. Part III draws out conclusions and implications from the literature and the findings.

Chapter 3 of this thesis describes the research methods which were used in the attempt to find answers to the above research questions. Like the whole thesis, Chapter 3 emphasises: that the methods were used in the context of a *less-industrialised country*; the *individual experiences*, including feelings and thoughts, which were an integral part of being a researcher; *practical* as well as theoretical aspects of the research methods.

All the other chapters of the thesis address the main research question of the roles of assessment in primary education. Chapter 2 provides some theoretical perspectives on, and practical examples of, official purposes of assessment and their political, educational and social effects. It draws on documentation from north and south but with an emphasis on the south where possible and on documentation about primary rather than secondary education where possible. Chapter 9 revisits these theoretical perspectives, but this time with reference to the results from the field research as well as with reference to the aims for the research which are outlined in this chapter, above.

Chapters 4 to 8 present an analysis of findings relating to the four subsidiary research questions (above). Chapter 4 describes official intentions for assessment in Egypt, both historically and currently, addressing the first research question. Chapter 5 describes how those intentions were implemented generally in the case study primary school, also addressing the first research question but at the same time throwing some light onto the issues in the second and third questions. Chapters 6, 7 and 8, respectively, present an analysis of pupils', parents' and teachers' positive and negative experiences of assessment and how they used it in their lives. They therefore address the latter three research questions.

## **Chapter 2. The roles of assessment in primary education: literature review**

### **2.1 Introduction**

Frederikson and Collins' (1989) words have already been quoted, that, 'The goal of assessment has to be, above all, to support the improvement of learning and teaching' (p.32). In this thesis, the assumption is examined that assessment is concerned with the improvement of learning and teaching, in some sense. A description of the kinds of improvement assessment is concerned with, provides an important means of defining distinct roles for assessment. This chapter is divided into three sections, each of which relates to a particular assessment role in terms of the official, stated purposes of assessment and the kind of improvement each is concerned with. Within each of these, assessment's official, stated use, its other uses and its obvious effects as well as subtler ones are analysed. The three major roles in terms of official stated purposes are as follows. Assessment may be for

- **the certification and selection of the individual**
- **accountability and system improvement**
- **the improvement of the individual's learning.**

### **2.2 The role of assessment for certification and selection**

#### **2.2.1 Assessment for certification and selection: its contribution to improvement**

Assessment for certification and selection stands apart from assessment for accountability, system improvement and improved individual learning in that assessment for certification and selection does not necessarily have *improvement* of educational achievement as an end in itself. Economic and political improvement may be the implicit aims of certifying and selecting the most talented.

Assessment for selection (though less so for certification) is about finding the highest achievers and not necessarily about improving lower achievement, although lower achievers may be inspired by selection examinations to improve their own achievement. Despite the increasing importance of and hopes in assessment for accountability and



system improvement, assessment for certification and selection purposes still dominates the world assessment scene (Little and Wolf, 1996). Lockheed wrote that World Bank support of assessment for selection (including certification) grew steadily since 1980 and consumed the largest portion of assessment funding (Lockheed, 1996). Heyneman and Ransom (1990) wrote that selection examinations were ‘... re-emerging as instruments of educational and occupational selection’ (p.181). Assessment for certification and selection has a dominant role to fulfil: the extent and nature of its contribution to improvement is therefore important. This issue, as well as a definition and description of assessment for certification and selection, is addressed in the following section.

### **2.2.2 The role of assessment for certification**

One of assessment’s roles is certification; this has been the case both traditionally and recently. *Certification* denotes a *declaration* or *formal statement* (Fowler and Fowler, 1978) which, in the assessment context, details specific educational performance. The purpose of certification is to provide an employer, an educational establishment or a child with a report about the child’s learning achievement. A certificate provides a summary of achievement which can be understood by others.

Using Berlak’s (1992) words, the purpose of assessment for certification is to

... serve as dependable national *educational currency* for individuals, that is to say, provide teachers, college admission officers, prospective employers, etc., as well [sic] parents and students with a record which accurately depicts a person’s strengths, what he or she knows and can do best (p.187).

Where the purpose of certification is to provide information to employers, its use at primary level is limited to countries where employment is an option at the end of the primary phase. In this case, the means of certification is usually an examination which serves to prove what the child could do at school. Where the purpose of certification is to inform an educational establishment, probably the junior secondary or preparatory level school, the means of certification may well combine with means for selection as an end of primary phase examination which predicts what the child will be able to achieve, on the basis of what she or he has achieved so far.

Where its purpose is an account for the child’s own record and motivation, certification may take the form of continuous assessments in primary school. Its purpose is still

summative, but it often summarises smaller units of achievement than certification for employers or educational selectors. In the first cases, certification often simply records examination results, although a more descriptive and individualised certificate might provide the employer or educational establishment with more useful information than an examination grade. Where certification is a broader record of ongoing achievement for the child's own use, more descriptive certificates in the form of records may be used.

### *Certification by examination*

In the few countries where employment is still an option at the end of the primary phase, such as Lesotho, Tanzania and Togo (Kellaghan and Greaney, 1992), an end of primary school examination score could provide potential employers with information about a child's skills. In reality, certification by final examination often plays an indistinguishable role from the selection role of assessment, since some form of post primary education is available in all countries. Describing the emergence of assessment for certification, Gipps (forthcoming) observed that certification became linked to assessment for selection only after an initial period when assessment had functioned for selection purposes alone. Today, the emphasis is still often directed in favour of predicting the future performance of the few pupils who are to be selected for further education, rather than on ensuring suitable employment for school leavers.

Kellaghan and Greaney's (1992) study of 14 African countries showed that all the countries held examinations at the end of the primary phase. They described a growing tendency in recent years to use the same test to play certification and selection roles. In Mauritius, for example, there used to be two examinations, one to certify which skills the child had mastered in primary school and the other to select children who would benefit from post primary education. Later,

... a realization that the two formal examinations contributed to excessive pressure on ten- and eleven-year olds led to the dropping of both examinations in favor of one (CPE) that now serves both functions (Kellaghan and Greaney, p.30).

Somerset (1987) described how the Kenyan certificate of primary education (KCPE) was reformed expressly to serve the needs of those primary school children who went directly from school to, often rural, employment, as well as to serve the selection purpose. The certificate at the end of the primary phase now assessed knowledge and skills relevant for the world of work. In this sense, certification was the reward for a

pupil's attendance at school which also proved that he or she had gained skills relevant to work in a rural community. In this way, children who failed to be selected to further education at least had a useful record of past achievement.

Oxenham (1984), however, showed how certification not associated with selection might be regarded as not important. He quoted the example from Mexico where many primary pupils in rural areas had no ambitions either towards modern sector employment or higher schooling. He noted that the teachers felt they should be teaching towards assessment for selection and so geared their curriculum around Spanish and mathematics; while the pupils, by contrast, seemed to attend school simply to gain some basic skills. In this sense, they certified themselves (Brooke and Oxenham, 1980). Oxenham assumed that these children were destined to work in familiar circumstances where employers needed no formal certification.

### ***Certification through alternatives to examinations***

When certification is for the benefit of employment or further education, both employers and higher educational establishments learn what the certificate signifies in terms of general curriculum coverage, often assuming that the examination is accepted as the most objective means by which to select children. Where that examination relates directly to a particular curriculum and that curriculum embodies skills, knowledge, understanding and attitudes relevant to employment, the examination result will be more informative, allowing employers and establishments to select appropriately.

Certification may be also for the benefit of the child's individual learning or the teacher's teaching of the individual, in which case it may provide more detailed information about actual skills, knowledge or understanding; unless it is just for the child's motivational purposes. Certification may occur throughout the primary years: mastered skills may be recorded as one school year ends and another is about to begin (Airasian and Madaus, 1983); certification may take the form of *continuous assessment* (Pennycuick, 1990), whereby many assessments, rather than one terminal assessment, are made throughout each term of the child's time at primary school. These assessments could be short tests of a particular curriculum area (Close and Brown, 1987); or they could be collections of the child's 'best' performance (Gipps, 1994), perhaps in an individual portfolio; or they could be Records of Achievement (Fairbairn,

1988) or in the form of tick lists; or teacher comments based on observation and discussion with the child; or they could be units of a commercial scheme recorded as finished (Sizmur *et al*, 1995). Wiggins (1989) described *exhibitions of mastery* (p.42) whereby pupils, albeit in his case secondary pupils, presented orally and visually what they knew and could do in various areas, which is a reminder that oral presentation at any phase is an alternative form of gaining certification.

### **2.2.3 The role of assessment for selection**

The purpose of assessment for selection is to make a summary from which to predict children's suitability for a higher class, a further phase of education or particular employment. On the basis of this prediction, some children are selected, others rejected. For example, in Barbados, an 11+ examination selected children for variously rated secondary schools; only children scoring the highest marks on the examination went to the best secondary schools (Payne and Barker, 1986). In Africa, the percentage of students who progress from the last grade of primary school to the first grade of secondary averages 43 per cent, depending on children's performance in an end of primary school examination (Kellaghan, 1992). Many have pointed to the negative direct and indirect effects that traditional selection examinations can have on children's learning because of their power to affect the child's future (for example, Alderson and Wall, 1993; Dore, 1976; Heyneman, 1987; Lewin, 1980, 1984; Little, 1984; Madaus, 1991; Meng Hong-wei, 1996; Payne and Barker, 1986). These effects are discussed in the last part of this section.

#### ***The purpose and means of assessment for selection***

By far the most common means of measuring children's suitability for promotion or employment is the formal examination at the end of a school year or phase, even though the predictive value of selection examinations is not proven to be good (Tolley, 1989) and the curriculum assessed in the examination is not always relevant to the next stage (Bracey, 1983; Dore, 1976; Kimberley *et al*, 1989). Formal, or psychometric, examinations are considered useful for selection since they allow comparisons between individuals; and *norm referencing* means that a predictable number of children are bound to do well and a predictable number to fail on any test (Gipps, 1994). A predictable number of places can therefore be filled at the next stage.

Aptitude tests have been supposed to preclude the possibility of excessive cramming and to reflect less cultural bias than content-based examinations, and so be a more acceptable and reliable means of selection. Yet, it has been found that measured IQ or aptitude is not constant, that its development may be closely related to factors in the social environment and that IQ tests '... confused cultural and environmental differences with innate properties' (Gipps, forthcoming). The 11+ aptitude test which selected children aged 11 was abolished in many parts of England partly because of its poor predictive value (Payne and Barker, 1986). In addition, ways of *cramming* for *aptitude* tests have been found (Little, 1984).

In some countries, for example China (Lewin, 1997), Pakistan (Hargreaves, 1997b), Kenya (Somerset, 1982), government schools may select pupils to enter the first grade of primary school by selection test. Payne and Barker (1986) described how in Barbados '... it is not uncommon to find pupils streamed from the reception class of the infant school' (p.314). Or selection may begin with promotion tests for entry from one infant class to the next; that is, where 'redoublement' is practised, which is especially prevalent in Francophone countries. The negative social and emotional effects of this practice have been clearly documented (Niklason, 1987) as has its economic inefficiency (Olivares, 1996). As selection mechanisms, though, promotion examinations effectively divide off those who are predicted to be able to cope with the next phase, from those who are not.

Although it is widely practised, Bowman (1970) questioned the whole assumption that selection by qualification was moral, or even logical, except in terms of using human capital most productively:

What, then, are the grounds for using control by qualification? One notion... is that people born with brains are somehow nobler, or 'more deserving' of special rewards. This is intellectual elitism in its most arrogant form. The second is a belief that the more able will benefit more from additional schooling *and* that individuals in making their choices will not allow adequately for that fact; this is a special version of the 'superior wisdom' theme. Third, it could be argued that the external benefits are especially great from educating the most able individuals. Only the third of these arguments, I submit, has even a whisper of empirical and moral validity (p.157).

### *Effects of selection in productivity*

Bowman (above) meant that selection was more useful for employers or governments who hoped to benefit from the advanced education of selected children, and less useful for the children themselves. Schooling may appear to serve individuals while, in reality, it functions as an initial input in anticipation of eventual economic gain for the state. If economic output is what policy-makers seek of education, it makes sense only to select those likely to produce returns. Heyneman and Ransom (1990) claimed that this was especially the case in developing countries, where *improvement* through education is seen more in terms of economic improvement. Lulat (1988), however, in his hilariously written review, disagreed with this assumption:

... what about all those numerous 'rates of return studies' much beloved by people such as George Psacharopoulos of the World Bank? ... all that these studies demonstrate is that some occupations carry with them higher incomes than others... The basic function of education, especially *general* education is to serve as a selection mechanism for employers; and it is a mechanism that has little to do with the productivity merits of the employee (p.320).

Lulat held that selection in education neither contributed to productive employment nor to the state or national economy; but rather promoted the enduring privileges of a social elite. Hallak (1990) agreed that the linkage between human resource development, or education, and economic progress, was by no mean direct. The 'human capital approach', which applies to education generally but especially to selection within education, has recently been questioned by educationalists and economists alike; Fagerlind and Saha (1989) described how

In the early 1960s the investment value of education for the improvement of human productivity was virtually unquestioned; by the end of the same decade human capital theory as the basis for a viable development strategy had been brought into doubt. No longer was it universally accepted that an increase in educational expenditure and of participation rates was sufficient to improve economic productivity both in industrialized and non-industrialized countries (pp.48-49).

Various forms of the human capital perspective are, none the less, still implicit within selection mechanisms in industrialised and less-industrialised nations.

If the role of selection at primary level, of a limited number of children to further education, is founded on economic scarcity, or the inability to provide enough places for all children, rather than for the benefit the individual or the child's future employer,

why does the individual child strive towards selection? Why do the parents of all children, selected and rejected, pay the same taxes towards education? The following section examines Lulat's suggestion that '... the basic function of 'general-education' is not economic, but rather sociological' (Lulat, 1988, p.319).

### *Effects of selection in social elitism*

Bourdieu and Passeron (1976) suggested that the middle classes, unable to perpetuate their status through capital alone, could fall back on a second line of defence: a school system which allowed them to perpetuate their privileged position by giving them a better chance of educational success. They described the notion of cultural capital in contrast to financial capital.

Foucault (1977) developed the concept of *normalisation*: rather than as the process of making normal or standard, he defined it in the sense of an authoritative standard (Copeland, 1996). Foucault's concept of normalisation involved the five processes of comparison, differentiation, hierarchisation, homo-genisation and exclusion. Meadmore (1993) described Queensland's primary State Scholarship Examination as the '... state aided education for a small, socially elite minority' (p.63). She used it as an example of the *normalising* function of assessment, whereby children learn to compare themselves with the norm and are judged according to their position in relation to the norm. That norm is represented by the small, socially elite minority who own the cultural capital.

Berlak (1992) pointed out that the norm transcended, and so in effect perpetuated, social context and history: the norm was imposed on children '... regardless of their families' history, culture, or race; regardless of gender...' (p.13). Children from the 'wrong' background fell outside the norm. Such children could be kept outside this 'norm', for example, because of the language of assessment, especially where a language is used which carries connotations of elitism. One example is the use of classical Arabic in Algeria and Tunisia (Heyneman and Ransom, 1990) or the use of a metropolitan language, such as French in Burundi (Eisemon, 1990). A politically biased language could further alienate those who speak other languages more fluently: Kellaghan (1992) showed how in post-colonial countries of Africa, the system itself focused resources and attention on students who could successfully negotiate the system and take their place in a colonial structure, putting others at a disadvantage.

Copeland (1996) described how modern science could justify how the norm was defined:

... the norm is the economy of social distribution, the process of objectifying subjects. The norm is supported by the work of the relevant sciences which, in turn, inform the dividing practices (p.382).

In an assessment system where being selected denotes elite status, those who are already elite attempt to keep their superiority; those who feel hopelessly excluded drop out of the race; but those who stand a chance, even a slim one, of becoming included instead of excluded, try to become so. In Dore's (1980) words, these are

... those who are bright, without being the brightest, those who are within sight of whatever are socially defined as the desirable prizes in the competition but by no means certain of reaching them without a great deal of anxious effort (Dore, p.55, 1980).

### ***Effects of selection in social mobility***

Wolf (1996) stated that:

... education in modern society: that is now the dominant route world wide for passing on elite status, as well as for determining life-chances and mobility generally (p.289).

Somerset (1996) described selection examinations as

... a second channel of elite recruitment: a channel based on 'merit' rather than on family privilege (p.265).

Examinations are seen as a useful tool for social mobility since, in Berlak's (1992) words, they represent *contest* mobility, through competition, rather than *ascribed* mobility through tradition, connections or family wealth (p.189). For a governing power, then, they appear to be a fair way for deciding who takes particular powerful positions. For individual learners and their parents, faith in qualifications becomes extreme, since they are seen as a powerful means that anyone can use to improve their lot.

The faith in the power of examinations for social mobility is especially strong in less-industrialised countries, where Little (1978) found expectations about future employment more unrealistic than in industrialised countries. In Ghana, Little (1984) found that belief in the power of hard work towards good school marks was inversely related to economic status. Seginer (1983) found a similar exaggeration of hopes



among lower class black parents in the United States of America and concluded that lower class parents relied exclusively on school grades since ‘... they have no other source on which to found their expectations’ (p.11).

Where children and parents see hope of betterment through selection, assessment for selection gains great importance in their lives. Poor examination results may deny students access to certain levels and forms of advanced schooling and thus close the doors to social, political, and economic advancement (Gipps, forthcoming). ‘Because written exams were associated with high-status professions, this type of exam became itself invested with high status’ (Gipps, forthcoming). Kellaghan and Greaney (1992) provided the text of an examination essay in which a Kenyan student described the fundamental importance to his life of his examination success:

There was suddenly a cry of joy and everybody came rushing towards me. It was so sudden that I almost fell off my seat. Everybody was hugging me and doing all sorts of funny things. I just could not believe my eyes. What I had been waiting for so anxiously was now revealed. It was too good to be true. I do not know how I made it, but passing my examination was the best news I had ever heard. None of us could control ourselves (p.80).

Attempts to change the system ‘... risk challenging the many vested individual and institutional interests in maintaining business-as-usual’ (Berlak, 1992, p.7). In other words, resistance may be expected. Reform was overturned, for example, in Sri Lanka when the potential for social mobility was threatened in the 1970s as British based General Certificate of Education certificates were replaced by local examinations (Little, 1996b).

Heyneman and Ransom (1990) also remarked on resistance to non *academic* examinations and curricula because of the association between elitism and non manual labour:

Parents aspire to ensure better lives for their children, usually equated with jobs in the modern sector that are better paid and less damaging to one’s health than is working with one’s hands. If the curriculum does not parallel those aspirations, parents react (p.189).

### ***The role of objectivity in assessment for selection***

Since selection examinations have such a vital role to play, and since they represent *contest* mobility, the contest must be fair. Faith in the mobilising power, but also the objectivity, of selection examinations can be startling. Objective examinations have

'highly replicable and reliable scoring' (Gipps, forthcoming) and therefore seem to be the fairest means of providing equal access to the opportunities they open up. Results which are objectively achieved cannot be disputed, since they favour neither rich nor poor. This belief follows world moves towards democratisation, away from individual favouritism and the power of wealth (Heyneman, 1987). Murphy and Torrance (1988) described how selection examinations were introduced into eighth century BC China precisely to preclude favouritism; but they added:

The tensions between the inequalities introduced by assessments and examinations and the inequalities which they seek to combat have been with us ever since (p.16).

While equal access through unbiased assessment is a constructive and logical ambition, many question its reality:

... tests are concerned with reality as constructed by culture. They are not, and cannot be, measures or indicators of some purely objective, independently existing state of affairs (Madaus on Hanson, 1994, p.222).

Gipps (1994) saw the notion of objectivity stemming from the 'science' of psychometric testing. She showed how the notion of objectivity could be used to justify the maintenance of an elite agenda. She wrote that performance

... is construed according to the perspectives and values of the assessor - whether it is the one who designs the assessment and marking scheme or the one who grades the open-ended performances. We do not therefore see assessment as a scientific, objective, activity, this we now understand to be spurious (p.167).

Bowman (1970) added that '... men [sic] want to be democratic and elitist at the same time' (p.147); that is, they want to appear democratic or objective, while actually perpetuating elitism. But while exaggerated belief in the objectivity of examinations has allowed the 'values of the assessor' and the system to dominate examinations unchallenged, alternative solutions towards equal access have not been found (although Oxenham, 1984, suggested '... it is more likely that non-awareness and inertia have impeded the search for alternatives than that such alternatives cannot be devised' p.14). Instead, the examinations themselves have become the focus of improvement (Heyneman, 1987; Little, 1984). They can become *more* objective, that is less biased towards any one group, if they are extensively trialled in negotiation with teachers, children and even parents (Gipps, 1994). Particular emphasis on informing and being informed by groups most likely to be targets of discrimination would help compensate for the dominance of one elite (Berlak, 1992).

### *Some negative effects of examinations*

Selection is most commonly determined on the basis of written examinations, although examinations may be used for accountability purposes too. Examinations may be high-stakes in terms of selection or of accountability or both. *High stakes* has come to mean how much the pupil, parent or teacher has to lose by poor performance in the examination. The extent to which stakes are high correlates to the extent of negative effect on pupils, parents and teachers. If they are high-stakes in terms of selection, in that a pupil's life chances depend on the results, negative effects will impinge mainly on the pupil, but also less directly on the pupil's parents and teachers. High-stakes examinations for accountability purposes will affect teachers more than pupils and parents, since it is teachers' professional reputations which are most at stake. In examinations for accountability purposes, teachers become accountable to *stake holders*, or those who would suffer (usually financially) if teachers did not produce the required outcomes. In the following section, negative effects on what is learnt and taught are considered, which may be similar for both types of high-stakes examinations; secondly, negative effects of high-stakes examinations on pupils, parents and teachers themselves are considered, although these may not apply to pupils and parents so much, when examinations are administered chiefly for accountability purposes.

Dore (1976) wrote that learning was sacrificed to a rat race for diplomas in countries where opportunities for higher schooling and employment were limited. In a system with *educational tunnel vision* (Lee and Ninnes, 1995), teachers teach and children study only what is assessed in the selection examination (Sislian, 1984). Of course, high stakes selection happens at certain stages only, for example at the end of primary school; its *backwash* or *washback* effects (Alderson and Wall, 1993), however, can be felt way down in the early years (Hargreaves, 1997a; Lewin, 1984). Meng Hong-wei (1996) wrote of the Chinese examination system:

[it] is called 'the baton' (*zhi hui bang*), directing the teaching and learning at which it points... schools try to achieve as high a pass rate as possible. This is reflected in the Chinese saying, 'thousands and thousands of soldiers marching across a single plank bridge'... The problem is that this march towards examination success violates the principle of 'all-round-development', which also underpins contemporary Chinese aims in education (p.157).

Madaus (1991) described how examination orientation could inhibit ‘... the development of the curricular variety that may be necessary to serve local and individual needs’ (Madaus, p.229); as well as encourage rote memorisation to the detriment of problem solving skills or creativity. Corbett and Wilson (1988) described how ‘... a tradition of past exams develops, which eventually de facto defines the curriculum’. Kimberley *et al* (1989) decried the need to divide knowledge and skills into subject domains for assessment, since this could fragment learning. Sizmur and Sainsbury (1997) complained that each subject that could be examined would be further fragmented into its assessable and non assessable components. Lewin (1984) illustrated how multiple choice formats especially encouraged fragmentation as well as rote memorisation, and limited emphasis on expressive writing. Bracey (1983) emphasised the limited skills that could be assessed in an examination: teaching towards these often meant teachers taught and pupils learned test taking skills, rather than the subject assessed.

Little (1984), among others, described the divisive social effects of pupils and parents feeling the need to buy expensive, therefore elitist, examination cramming, including private tuition, and the long hours they might spend in taking tuition. Teachers sometimes gave more emphasis to providing this than to classroom teaching, because it was better paid and conditions were more comfortable. Smith (1991) described the implementation of the Iowa Test of Basic Skills for accountability purposes, as a ‘cruel and unusual punishment’ for young children. She stressed that ‘... many teachers themselves feel anxious, worrying about whether they have adequately prepared their pupils for the test’ (p.9). She went on to say that

Faced with the ‘packed curriculum’ (a set of requirements... that exceeds the ability and time of any teacher to cover all of them competently) and the restricted number of instructional hours available, some teachers... will suffer sanctions and loss of autonomy because of low scores (p.10).

Madaus (1991) described how examination orientation could encourage emphasis on scores, rather than diagnoses of weaknesses in both learning and teaching. Little (1984) showed that pupils became more extrinsically than intrinsically motivated to learn, when their focus was good examination marks, and teachers accordingly used extrinsic motivations to help them learn. Bernstein referred to an openness and closedness of

contents, the closed which were collected to satisfy evaluation criteria, the open which related to each other:

The educational identity of the student is shaped by the isolation of contents:... [and] evokes a gradual realization in the student of becoming different from others and of belonging to a particular category or even self-conscious group (Cherkaoui, 1977, p.559).

That is, pupils form an identity on the basis of others' judgements which, at school, may relate almost entirely to their performance on tests: Hanson described how '... test scores transform people by assigning them to various categories in which they are treated, act and come to think of themselves' (in Madaus, 1994, p.232). When the focus of such judgements is on comparison or competition with others, children's evaluations of their ability and feelings towards themselves will be more negative (Ames, 1992). Little (1985) pointed out that, through the use of tests and studying for them, good and bad behaviour became determinants of success and failure.

In addition, Hanson (1993) warned that tests made pupils feel powerless, since they had no control over the tests and so became transformed from 'autonomous subjects to passive objects'; thus tests acted '... to curtail individual freedom and dignity' (described in Madaus, 1994, p.229, 223).

### ***Improvement and assessment for certification and selection through improved examinations***

The purpose of assessment for certification and selection is certification and selection and improvement in broad terms, such as a country's economic or political improvement. While this continues to be the case, some have argued that the assessment instruments used to certify and select could play a role in improving education directly, mainly in terms of what is learnt and taught. How high the stakes are will relate to how much impact the instruments may effect on learning and teaching. In this sense, the higher the stakes the better. In other words, by improving assessments themselves, teaching and learning will themselves be improved. Kellaghan and Greaney (1992) claimed for the purpose of their study of 14 African countries that

... by focussing on the close interrelationship between formal assessment, teaching and learning, it helps to pinpoint the way to raising the level and the quality of education of pupils in Sub-Saharan Africa (p.v).

Somerset (1987) described the system in Kenya where feedback on the results of high-stakes selection examinations reached all schools. Because the examination was so high-stakes or important, children and teachers directed themselves to the beneficial goals which had become part of the selection mechanism. For example, where practical or other relevant skills and those dependent on reasoning, problem solving or creativity were necessary to pass the selection hurdle, children and teachers quickly addressed themselves to these aims. Frederikson (1984) described how ‘... dramatic changes in achievement came about solely through a change in the tests’ (p.201). Heyneman (1987) described examination feedback mechanisms as ‘... the most efficacious means that educational managers have to improve the quality of education’; Popham (1987), coined the phrase ‘measurement-driven instruction’ and promoted it as a means to bring about improvement in teaching and learning; Wise (1978) discussed the power of minimal competency testing to raise specified standards.

The causal link between the nature of the testing instrument and improved learning and teaching is, however, not unanimously agreed, just as the exact relationship between assessment and its negative impact cannot be defined. While some perceive even good tests to have predominantly damaging effects (Berlak, 1992; Raven, 1992), others consider the relationship to be more complicated. Alderson and Wall (1993) described the belief in tests as ‘levers for change’ as a ‘naive deterministic view’ (p.118). They claimed that virtually all researchers who ‘proved’ that good tests could improve learning ‘... use anecdote, assertion, or interviews and surveys of what teachers and pupils say they do rather than direct observation’ (p.123). Dove (1986), provided one more aspect of this issue:

Tests which are technically and educationally sound can do much to improve the quality of teaching and learning. But this alone is not enough. The most important ingredient is the training of teachers to understand why their work in testing and assessing pupils is an integral part of their core educational role and training them in sound assessment techniques (p.69).

## **2.3 The role of assessment for accountability and system improvement**

### **2.3.1 Accountability in relation to improvement**

Assessment for accountability and system improvement has recently increased in popularity, especially in industrialised countries such as England and the United States of America (Little, 1996b). Wise's amusing article of 1978 cited 14 terms being used in the United States of America as part of the 'accountability movement', including references to planning, budgeting, management by objectives, operations analysis, management science and economic analysis. Just as certificates could be used as educational currency for individual employers or future educators (see above), so accountability measures could be currency for the suppliers and maintainers of the education system itself, in any country.

Lockheed (1996) showed that projects which aimed to monitor educational standards received the greatest increase in all World Bank funding for assessment since 1988. This increase was especially notable in Latin America and Europe, the Middle East and North Africa and applied more to basic education than any other phase. The increase may be accounted for to a large extent by the emphasis world-wide on quality as well as on quantity: the aims inspired by the Jomtien Education for All conference in 1990 included universal enrolment by the year 2000; but enrolment is no guarantee for quality learning. By making learning sites accountable to the government or organisation who supplies and maintains the system of learning sites, that government or organisation can check up that their recommendations for learning and teaching are being implemented or that money provided is being used appropriately. Assessing can be for financial accountability, accounting for the relationship between financial input and educational output; or it may be for political accountability, accounting for implementation of policy.

Indirectly, measures for accountability enable improvement, in that when a school is found not to be meeting required standards, either in terms of value for money or in terms of implementing policy, intervention takes place to improve standards up to the required level. By *checking up* on the whole system of education, the state gains guidance on the *building up* of the system. For this reason, assessment for

accountability is categorised in this thesis along with assessment for system improvement.

*Improving* education means making the education system more effective in relation to its stated aims; an improved system provides more appropriate education (Patrick, 1996) in relation to the system's aims. Improvement could involve making already existing provision more appropriate (Watson, 1984; Lockheed, 1996) or introducing an entirely new approach, as in South Africa where the previous education system of the apartheid era was thrown out in its entirety (CEPD/EPU, 1998). Olivares (1996), Somerset (1987) and Wiggins (1989) showed how an analysis of test results gathered during monitoring could direct action for curriculum and assessment improvement. The focus of improvement inspired by national monitoring data is improvement of the system as a whole, rather than the learning or teaching of individuals, and it is for this reason particularly that government makes accountability and system improvement official purposes for assessment.

### **2.3.2 Monitoring**

Monitoring projects are often set up specifically to ensure that schools demonstrate their accountability to a central controlling body, by submitting specific, standard data. The data collected in monitoring projects may describe various outcomes from schools, in particular, pupil attainment outcomes, which are generally reported in summary form.

Little (1996b) gave four reasons for undertaking monitoring:

1. For financial accountability. This may be especially important where international aid agencies provide resources for specific educational activities and wish to make sure the resources are being used as designed. Lewin (1997) described how, in China, extra resources were allocated to schools measured to be achieving most highly.
2. To detect time trends in achievement. In particular, governments monitor to check up that improvement rather than deterioration or stagnation is coming about. Somerset (1996) showed how annual *league tables* of examination results made schools accountable and so encouraged primary schools in Kenya to improve instruction over a number of years. Wood and Power (1984) commented that '... the whole [standards]



debate has not been about what children can do... but about whether they are getting better or worse at doing it' (p.309).

3. To determine the causes of differences in academic achievement. Attempts can then be made to counter negative differences, sometimes using rewards and penalties. The UNESCO-UNICEF monitoring project tried, for example, to link learning achievement to the characteristics of the child, home background factors, the school setting, teaching-learning factors, the community environment and selected school-community factors (Chinapah, 1996; Meng Hong-wei, 1996). In Korea, from 1983 to 1985, a longitudinal study of problems with the elementary curriculum was made in order for reforms to be made to the curriculum (Chan-Jon Kim, 1996). Chan-Jon Kim described how poor test results were sometimes used in Korea to reprehend those in charge of the districts where they occurred.
4. To evaluate the effectiveness or efficiency of educational innovations. Where a new system has recently been introduced, such as the new national curriculum in England in 1988 (Kimberley *et al*, 1989) or South Africa in 1998 (CEPD/EPU, 1998), government wishes to demonstrate its success.

The term *monitoring* can be used to describe a wide range of purposes and methods. Monitoring involves some form of regular surveillance, involving the measurement of, or gathering of information about, the performance of the educational system. The means used to monitor may include quantitative and qualitative data collection methods, but surveys of pupil outcomes are a common means. The means include standardised or other tests or examinations taken by children specifically for the monitoring exercise; the collection of background data; inspection visits and reports; questionnaire surveys about practices or attitudes; and occasionally, interviewing and observation. Monitoring is the main means of assessment for accountability, although results of assessments for certification or selection may also be used for monitoring purposes.

Whole national cohorts may be monitored, for example, all fourth year primary school children and eighth year pupils were monitored in the Chilean Quality of Education Assessment System (Olivares, 1996); or a sample of children may be monitored, as in the Chinese study on moral education where nearly 1500 grade 5 children were chosen



across 11 areas of China (Meng Hong-wei, 1996); monitoring exercises for accountability and system improvement often include nationally representative samples of schools or pupils. One or two curriculum subjects may be monitored, as in the 1993 to 1998 International Association for the Evaluation of Educational Achievement (IEA) study of science and mathematics with nine and thirteen year olds (Goldstein, 1996); or broader aspects of achievement may be monitored, such as in the UNESCO-UNICEF monitoring project of literacy, numeracy, life skills, factors influencing learning achievement and equity issues in basic education (Chinapah, 1996). Monitoring may be carried out within one local area or across many countries. Gipps (forthcoming) commented that

A fascination with performance indicators and international comparisons of performance encourages governments not only to use assessment to drive and control their curricula and teaching, but also to invest large sums of money in international surveys of educational achievement.

### **2.3.3 Intentions for system improvement and their realisation**

Wood and Power (1984) doubted the simplicity of using monitoring data to inform system improvement:

... the implications seem to be that everyone will know what to do when they receive the monitoring results. They do not, of course (p.316).

Wolf (1996) likewise pointed out that designing and legislating policy for improvement did not mean '... it will somehow self-implement' (p.287); Hallak (1990) stressed that diagnosing a need for resources during a monitoring exercise and then providing them did not mean the resources themselves produce desired educational outcomes; Gipps (forthcoming) suggested that testing programmes could be symbols of order and control, linked to traditional moral values, and for this reason rather than from empirical evidence of their likely effectiveness, they struck a '... responsive chord with the public at large' (p.11); and Madaus (1985) saw testing as a convenient way for policy-makers to side-step the problem of dealing directly with the instructional process, thus transforming testing into a coercive exercise.

When monitoring, assessment tools have to be used which best allow an analysis to be made about the current state of the education system. Priorities have to be made as to which indicators to use and then assessment instruments chosen which best reflect the

indicated aspects of the system. Wood and Power (1984) expressed the difficulty of relying on monitoring data to inform appropriate improving action:

Any attempt to assess national educational performance must reckon with the instrument getting in the way, to the extent that ultimately the results and the characteristics of the instrument which produces them are totally confounded (p.313).

The suggestions made here are that the relationship between monitoring for accountability purposes and ensuing intervention for improvement in education is complex. Lockheed (1996), in relation to World Bank support, described how the central authorities could potentially exert extensive control through assessment for accountability, control which had only partly to do with educational performance *per se*. While claiming to be maintaining and improving educational *standards* through assessment for accountability, central authorities may not be willing or able to articulate more far-reaching improvements they hope to make: for example, improved social cohesion, through the transmission of a set of standard skills, facts and values; economic improvement through a better-skilled workforce; contentment from national stake-holders, as well as observers from other countries, through the use of nationally and internationally acceptable programmes. For example, Lolwana (1996) wrote of the *punitive* state system of South Africa before 1994, which legislated racism through education and its assessment. Moving to industrialised England, Hextall (1989) related assessment policy to the then prime minister in these terms:

... the current proposals on national assessment/testing cannot be disentangled from other aspects of the Thatcherite programme of educational restructuring... the very context, structure and process of such modes of assessment are inherently (perchance, intentionally) oppressive and discriminatory (p.239).

Lulat (1988) described how the state's foremost interest could be its own preservation:

... there is a notion that the state exists on behalf of society as a whole, and consequently as a representative of common interests of society... with respect to the circumstances of most TWNs [third world nations] the ridiculousness of this notion is self-evident from even a cursory examination of what goes on among them:... often outright robbery of the public treasury... (p.324).

Less-industrialised nations may inherit highly centralised systems from colonial powers, especially former French colonies (Watson, 1984). Using assessment for accountability, newly independent governments may strengthen this inherited centralised power, using educational assessment as the tool to achieving extended power in broader social,

political and economic fields. Yoloje (1985) made the point that, through international funding, a centralised control could be imposed at international level, making national governments feel more need to exert their own power within their country. In addition, continuity, and so deeply entrenched power, may be assisted by international accountability systems.

State control of the system, however, does not inevitably mean the abuse of a strong centralised power. On the contrary, strong central control of education through assessment for accountability and can foster a powerful sense of national identity and cohesion (Watson, 1984). In addition, a strong centralised power may tolerate various degrees of regional or local control. For example, in Indonesia, only 80 per cent of the curriculum is centrally prescribed to allow for local initiative (Umar, 1996).

#### **2.3.4 The effects on teachers, of monitoring for accountability and system improvement**

Once monitoring data has been collected, for example by a national Ministry of Education [MoE] or its agency, the Ministry must decide what steps to take on the basis of the information they provide. The results of data from assessments administered during a monitoring exercise may tell the Ministry:

1. That the performance it requires is being achieved or exceeded. This pleasing state of affairs may result from making learning sites prove their performance through the use of assessments; that is, because they have been motivated to reach specific standards, on pain of penalty, they work until they reach these.
2. That in some (or all) learning sites, the performance it requires is not being achieved. Improvement is therefore necessary in these sites in order that the whole system can be said to reflect certain minimum standards.
3. That where improvement is necessary, mechanisms must be decided upon and put in place to bring improvement about. Teachers will be implicated in mediating this improvement.

Teachers are affected by knowing that a greater, external authority is prescribing to them that they carry out assessments and that they submit the results of these assessments. Teachers are affected by the fact that they have no choice in this.

Teachers are also affected by realising that the results they submit must meet approval, otherwise they may be told to make changes in the way they work, either as part of individual remedial measures or of a national improvement programme, based on data collected during a monitoring exercise. The approval they need may not represent teachers' own values but, again, they have no choice.

The individual teacher is affected, more than the individual pupil, where the emphasis is on accountability rather than selection. Because measures for accountability are usually *summaries* of pupil attainments, it is the teacher rather than the individual pupil who is directly accountable to the monitoring authorities, albeit through the institution of the school (Brooke and Oxenham, 1984). She may feel she has a responsibility to individual pupils and, more likely, she may feel accountable to parents for individual pupils, and additionally, she may depend on parents' financial support in terms of salary or payment for private tuition. But she is accountable to the monitoring authority for *all* her pupils and for her capacity as a teacher. Schaeffer (1992) noted that '... personnel are accountable solely to superiors within the bureaucracy and not to the so-called beneficiaries' (p.16).

As an effect of externally imposed monitoring exercises, teachers change their behaviour in certain ways, because they must focus on the *reporting* requirements of carrying out assessments. With assessments for certification and selection, the emphasis is on the individual achievement of pupil and teacher to some extent; with assessments for formative purposes, the emphasis is on the processes of the individual pupil's learning and the individual teacher's teaching; but with assessments for accountability, the emphasis tends to be on the *summary of results* which are reported and then used externally.

Smith (1991) wrote of the anxiety associated with the pressure of assessment for accountability:

To understand the perceived effects of external testing on teachers, one needs only to ask. Their statements on questionnaires, in interviews, and during conversations in meetings ... reveal the anxiety, shame, loss of esteem, and alienation they experience from publication and use of test scores (p.8).

Dove (1986), on the other hand, described how the 'ritualism and heightened anxiety' associated with externally prescribed tests gives them something of the character of

*rites de passage*, so that those teachers who survive the pressure and whose results are approved of, feel they have really earned their success. However, the teacher may choose to ignore children's weaknesses and even to enhance her own pupils' results, because the emphasis in the monitoring exercise tends to be on results. Olivares (1996) described how teachers responded to SIMCE in Chile, because they perceived assessment results as an important teaching goal:

... they have tried to 'help' their pupils in unorthodox ways... there are heads of establishments who do not pass on their poor results to governors... (p.133).

The individual teacher's professional judgement may be undermined since to monitor performance, standards of good performance must be established by the monitoring agency rather than by teachers themselves. Smith (1991) commented on how this could stifle local initiatives. She wrote that

... the districts' drive to keep high scores high and eliminate low scores will prevent them from experimenting with promising programmes that are not closely matched to test contents (p.10).

Accountability exercises do not always include information gathering and dissemination actually within school. Reynolds and Cuttance (1992) stressed that monitoring information should be built up in collaboration with school departments and reported back to school departments if teaching and learning were to benefit, rather than just the outcomes of formal assessments. If teachers were not included in the monitoring process, Torrance (1984) held that an account given by the school would be rhetorical rather than useful and would not therefore reflect everyday classroom learning and teaching. Ensuing *system improvement* would therefore be a superficial improvement and teachers' professionalism would be undermined rather than enhanced.

State or national monitoring agencies may claim to consult teachers when deciding what constitutes good performance, but examples of very superficial consultation are common (Chan-Jon Kim, 1996; Kimberley *et al*, 1989; Olivares, 1996). Porter (1989) showed that curriculum standards imposed through power but which lacked authority in teachers' eyes, would be implemented reluctantly and so be unpersuasive and short lived in their effect. Shaeffer (1992) described how government officials often went directly from college into government positions. When checking up for accountability purposes, they had little appreciation of (or even scorn for) local or indigenous

knowledge or creative potential of what are seen as largely passive *beneficiaries*. This top-down mentality could

... lead to the discouragement and even the fear of diversity, local initiative, and conflict; the avoidance of open consultation and interaction, except with other professionals... the desire to control information, events and people; and, ultimately, a resistance to change in structures or procedures (p.15).

Schaeffer therefore stressed the importance of participation and consultation in any national educational initiative if teachers' professionalism were to be enhanced, and the necessity for a conduit for information from the school upwards as well as from the centre down to schools. Black (1994) described teachers after the imposition of the national curriculum in England:

There was much anger and demoralization among teachers, who feel that nobody understood or cared about their needs and that national curriculum and assessment have become political footballs. They have learned that the mass boycott is the only effective weapon (p.199).

Demoralisation among teachers may result if teachers feel their professionalism is not respected. They may then respond to top-down prescription with apathy, considering the penalties of expressing an alternative opinion too great. This applies especially in the many less-industrialised countries where teachers have no initial teacher training. Dove (1986) wrote:

... if teachers have low levels of general education and are untrained or inadequately trained, they are likely to lack confidence and ability in undertaking any activities other than those which involve familiar and 'safe' teaching routines (p.59).

However, Dove (1986) described how in Malaysia, '... not only is it not right for the young to question their elders but teachers also share the same conservative attitude and are used to being directed from above' (p.49).

Finally, Avalos (1990) suggested that teachers in less-industrialised countries, as a group, enjoyed less overall autonomy in the making of instructional and assessment decisions with regard to their classes, than in more industrialised countries and therefore external prescription may be important as a support to teachers. Oxenham (1984), comparing two very different improvement situations, those of Ghana and Mexico respectively, described parallel reform efforts. Both involved all primary schools nationally, and both aimed towards increased 'intellectual stimulus, local relevance and co-operation' in primary education (p.158). Ghana had a centralised

examination system which subsequently little moulded itself towards the new reforms; the teachers therefore continued to teach to the centralised test since pupils, parents and teachers alike seemed happy enough with the institution. Mexico reformed assessment by leaving its design to individual teachers, but provided centrally developed support material, but teachers there used this new flexibility to teach in the least tiring way possible.

### **2.3.5 The effects of assessment for accountability and system improvement, on pupils**

The preceding sections highlight the possibility that, when assessment for accountability and system improvement is imposed from the centre, teachers may only make as much effort as benefit their results, rather than entering the spirit of the prescription for improvement. Their own feelings of anxiety and alienation affect the way they treat pupils. Little (1990) observed that what is facilitative to one party, such as the teacher, may be inhibitive to another party, such as the individual child. Monitoring is not necessarily for the benefit of the individual child's learning because educational improvement is seen in national rather than individual terms. Wolf (1996) showed that, upping the *stakes* for teachers, schools, districts or nations still did not make an assessment important, or indeed useful, for the child. The purpose of monitoring for accountability is accountability to parents on a micro level and to the suppliers and maintainers of the whole system at the macro level.

Wood and Power (1984) suggested that,

The sort of bold alternative which would attempt to associate 'standards' with, say, a developmental view, in which the growth of meaning, understanding and intellectual processes are important, and the fostering of self-actualisation and autonomy imperative, finds no mention. Standards, according to this paradigm, would be represented as complex descriptions of process and stage of development reached - a totally different formulation from a series of facility values on largely unconnected items (p.319).

That is, parents or stake-holders may be steered to concentrate on facility values, or a grade or mark, on the basis of a written test in a particular subject area which only give them some relevant information. Their interpretation of that information depends on how well-versed they are in what the grade or mark means in terms of certification or curriculum coverage. Parents may develop *examination tunnel-vision* because of the



way national data are reported. Some of the resistance to national monitoring in England and Wales came about because teachers misconceived the purpose of national assessment to be the learning of the individual, rather than school and school district accountability. Some teachers in England and Wales were therefore surprised not to 'learn anything new' about children's achievement while administering the assessments (Sainsbury *et al*, 1992).

In the United States of America, monitoring is most commonly carried out by testing children on standardised tests (Madaus and Raczek, 1996). Airasian and Madaus (1983) pointed out that tests usually tailored for measuring individual differences were applied unadapted to monitoring school differences; additionally, that standardised test scores used to measure school achievement hid unique and statistically significant school and programme achievement differences. In other words, monitoring was irrelevant at the individual level and inefficient at the school level, when standardised tests were used. School results, further, as comparisons with other schools, showed rank order but gave no indication of the kind of child within the school (Kimberley *et al*, 1989).

Somerset, however, gave a unique example of how the primary assessment system in Kenya combined a nationally prescribed accountability exercise with an analysis of pupil skills and knowledge, albeit at a general level, through the distribution of a newsletter (Somerset, 1987). In 1995, the School Curriculum and Assessment Authority in London commissioned error analyses of the national monitoring tests taken by seven year olds in England and Wales, but as yet these have only been distributed to teachers in abridged form (Hargreaves, 1995).

## **2.4 The role of assessment for the improvement of the individual's learning**

### **2.4.1 Assessment for the improvement of the individual's learning as an official, stated purpose**

In England, the national curriculum Task Group on Assessment and Testing (Black, 1988) recommended that national assessment served four purposes: formative, diagnostic, summative and evaluative. All four purposes were official purposes for a new assessment system. The terms have been and still are often confused, since the four terms are not equivalent. Broadly speaking, summative purposes include reporting assessment outcomes for accountability or for certification and selection; evaluative purposes include reporting assessment outcomes for accountability and system improvement; while formative and diagnostic purposes include improving individual children's learning, for example, in any one classroom. There is much overlap among the four types. For example, the results of assessments carried out for summative or reporting purposes may subsequently be used formatively, to improve the learning and teaching of individuals within a class (Gipps, 1996).

However, what is notable about the TGAT model is that diagnostic and formative purposes were separated from and given equal emphasis to summative and evaluative purposes, as official purposes for national assessment. Although subsequent events showed that this emphasis shifted more over to the summative and evaluative purposes later (Black, 1993), here was an example of the three official purposes for assessment discussed in this thesis, all being stressed by policy-makers. Prior to 1988 there had been no official policy on assessment for any purpose in primary schools in England, but

... the role of assessment in learning has in one sense always existed since it is difficult to conceive of a learning process which is not concerned with monitoring its own progress (Goldstein and Lewis, 1996, p.3).

The role of assessment for the improvement of individual or classroom learning was, perhaps, all the more important before TGAT *because* it was not dominated by national policy of assessment for other purposes. That is, assessment for the improvement of classroom learning may thrive even if it is not stated as an official purpose for assessment, or perhaps *because* it is not stated as an official purpose.

### 2.4.2 Formative and summative purposes

Assessment for the improvement of individual learning may be defined as assessment whose purpose is improved learning within the classroom or learning site, regardless of requirements by local or national authorities for the submission of assessment results. Its purpose is formative, that is, the teacher uses assessments which will provide her with information most likely to help her teach each child or a group of children more effectively (Black and Wiliam, 1998). Although her goal may be to improve assessment outcomes that will subsequently be submitted to local or national authorities, she does not submit the results of the assessments she uses for formative purposes directly, since their purpose is to be fed back into her teaching (and perhaps to parents). Broadfoot and Osborn (1987) noted how attention specifically to the individual rather than individuals within the group was not always considered to be necessary for learning; but in the context of this thesis, assessment for the improvement of the learning of the individual means that the individual child *or classroom* is the beneficiary of assessment rather than an outside authority.

Assessment for formative purposes is ‘... the process of appraising, judging or evaluating students’ work or performance, and using this to shape and improve their competence’ (Gipps, forthcoming). The purpose of such assessment is to help rather than sentence the individual, to support rather than to measure learning. Phrased differently, assessment for formative purposes refers to assessment activities

... which provide information to be used as feedback to modify teaching and learning activities in which they are engaged. *Such assessment becomes ‘formative assessment’ when the evidence is actually used to adapt the teaching work to meet the needs* (Black and Wiliam, 1998, p.2).

Evidence which is gathered during assessment for formative purposes will take the form of details, whether written or mental, about the nature of children’s learning. These details might relate to the child’s attitudes and approaches to learning, learning contexts which suit the child best, the depth and breadth of skills, knowledge and understandings the child owns and diagnoses of strengths and weaknesses in her or his learning. The child’s teacher, perhaps in discussion with the child, uses any evidence she considers relevant, in order to decide on learning and teaching strategies which will most

effectively improve the child's learning (Frith and Macintosh, 1984; Harlen and James, 1997).

Results of assessments for summative purposes, on the one hand, which indicate how effectively the education system is working, for accountability, are reported in summative form. This kind of evidence often appears as marks or grades, achieved by whole cohorts of children in specific, representative, curriculum areas such as literacy, numeracy and life skills (Chinapah, 1996). These summaries are reported to government or stake-holders who use them to make judgements in relation to maintaining or improving the whole system (see above).

In the summative case, then, the improvement focus is on outcomes of learning as indicators of the effectiveness of the whole system. In the formative case, the focus is on any aspects of learning which provide indications of appropriate next steps to improve the learning and teaching of the individual child. Because this kind of assessment is often practised through continuous, informal classroom assessment, rather than through a nationally prescribed assessment system, firstly, a formative purpose is less likely to be stated as an official purpose for assessment; secondly, many variations in the form it takes are to be expected. These variations are described below.

#### **2.4.3 Individual learning and assessment which helps improve it**

How evidence of learning is used, and which evidence is used, to adapt teaching to improve the child's learning, depends on the kind of learning the assessor believes should be, or is, taking place. The same forms or types of assessment may be used by assessors who hold opposing beliefs about learning, but they will use them to improve learning in different ways. In any case, their purpose is to find out how to improve learning among the children they are assessing.

The word *learning* itself may be defined variously as 'getting knowledge'; 'getting skill'; 'finding out'; 'committing to memory'; simply 'receiving instruction'; a mixture of these; or all of these together (Fowler and Fowler, 1978). Teachers and assessors do not necessarily hold consistent or constant beliefs about learning; many may subconsciously adhere to several and even contradictory perspectives on learning, which may depend on the content and context within which learning takes place. In addition, theory about learning is still unformed, with the exception of some areas of

language learning theory. And in many countries, including less-industrialised ones, the learning model may be defined more by tradition rather than by cognitive theory (Watson, 1984).

It could be said that there are two main sides to the argument about how learning takes place and therefore how assessment should be practised. On the one side, learning is seen to come principally from outside the learner and be added to the learner; it is not therefore dependent on the individual who is learning. On the other side, learning is seen to come from inside the learner, although this inner learning may be dependent on outside factors too. Within the two sides, learning is described in different ways. Whether learning is perceived more as an external addition or as an internal realisation, determines what form and type of assessment of such learning will be appropriate for formative purposes. By clarifying the perspectives, the most appropriate means of assessment and types of assessment information may be reflected on, which will enable the most effective steps to be taken towards improving each child's learning. The two perspectives and the epistemologies they reflect may be broadly described in the following different pairs of ways, although in reality such distinctions may not be at all clear.

<b>Learning comes from outside</b>	<b>Learning starts inside the learner</b>
Learning occurs by changing behaviour (behaviourist model)	Learning occurs through the use of innate cognitive faculties (cognitivist model)
Learning occurs when teachers transmit information (transmission model)	Learning occurs when individuals make sense of their experiences (constructivist model); learning occurs through social interaction (sociocultural model)
Learning is a surface-level activity	Learning is a deep-level activity
There are certain basic skills and facts which need to be learnt	Learners who have higher order skills acquire basic skills and facts for themselves

### *The behaviourist and the cognitivist models of learning*

In the behaviourist model of, for example, B.F. Skinner, learning was seen as behaviour change. Behaviour could be shaped almost at will by the teacher providing stimulus, the pupil responding, and the pupil then suffering the positive or negative consequences (OU, 1991). Learning was seen to occur in a linear series of tiny steps, each of which was subordinate to the next and had to be mastered before the next could be mastered. The teacher's task from this perspective was to provide the learner with a stimulus to which the learner responded and the teacher then gave positive or negative reinforcement to the learner's response. From such a perspective, assessment with formative effects would involve the assessor in recognising whether or not the learner responded appropriately to one stimulus and so was ready to receive another.

In Sadler's (1989) words, the teacher seeks '... information about the gap between actual and reference levels and uses this information to alter the gap' (p.121). On the basis of the teacher's diagnosis through assessment, she makes a formative judgement about the child's next learning objectives. Her formative purposes will determine the depth of her diagnosis, which could be a general analysis of which outcomes the child has or has not mastered or an analysis of exactly why a child has not mastered a certain skill. Each type of diagnosis or analysis will form the teacher's teaching in a different way. For example, she may decide to change her teaching plan for the following week since she sees that several children have still not mastered a particular skill or she may decide to give individual children remedial learning programmes to address particular difficulties. Such an approach does not belong only to a behaviourist perspective on learning, but when the teacher sees the objective of learning as the child's mastery of a particular and observable skill which she teaches, the teacher could be said to hold behaviourist views on learning.

Criterion referenced assessment for learning may reflect a behaviourist model of learning, if the learning and assessment criteria are considered discrete and hierarchical, and depending on the nature of the criteria. Through criterion referencing, the teacher provides objectives for the individual, rather than negotiating objectives with the individual. Criterion referenced assessment could take the form of 'graded tests' which, as Brown (1989) described, were '... generally organised into a set of successive levels'

which ‘... incorporated at each level a set of objectives (grade criteria)’ (p.300); or items in tests which assessed against a particular criterion of achievement (Pennycuick, 1988); or broader ‘graded assessments’ which were not limited to formal tests. If the child showed mastery against a criterion the teacher (or pupil) could make a formative decision about which criterion the child may address next. The pupil may receive certification of attainment which will help motivate the child’s learning.

Examples of the use of criterion referenced assessment which reflect some aspects of a behaviourist perspective on learning, are found in South Africa’s Outcomes Based Education (CEPD/EPU, 1997) and in England’s national curriculum (especially its 1989 and 1991 versions) and its assessments, although both of these include some elements of norm-referencing alongside criterion referencing.

Criterion referenced assessment, like the stimulus and response learning model, runs the danger of fragmenting information; that is, related ideas within a construct, or whole subject, are sometimes assessed in isolation so that the teacher may not know whether the child grasps the whole picture (Sizmur and Sainsbury, 1997). The fragmentation of ideas makes them more difficult for the learner to relate to situations outside the learning environment. Where criteria are presented hierarchically, also, it cannot be assumed that all children will learn in accordance with this hierarchy. In addition, knowledge is seen as an external body of categories which may or may not strike a chord with the learner’s perceptions from experience. But fixed hierarchical criteria are easy to use in teaching large numbers of pupils.

Behaviourist theories of learning, whereby learning comes from outside, have been challenged by cognitivist views. Chomsky (1980) reflected a cognitivist approach to learning when he wrote that ‘We may usefully think of the language faculty, the number faculty and others as ‘mental organs’’ (p.39). Gardner (1985) similarly postulated innate ‘intelligences’. Neither assumed that external factors did not influence learning, but that intelligence existed in potential form inside the pupil’s mind. Teaching, then, was about providing experiences which would enable the individual to realise her or his full cognitive potential. Such a view means learning must, to some extent, be individually tailored.

Traditionally, cognitive ability or intelligence has been assessed using psychometric tests.

But the traditional psychometric testing model was essentially one of limitation: the aim was to measure attributes which were a property of the individual and which were thought to be fixed (Gipps, forthcoming).

Assessment to encourage rather than to limit learning, as perceived from the cognitivist perspective, would provide the child with the context and experience within which to demonstrate full cognitive potential. Denvir (1989) proposed that each child should be assessed '... in a range of contexts and the contexts in which they are most likely to succeed should be sought, in order to accurately reflect mathematical understanding' (p.282). From a cognitivist perspective, errors in a child's performance are formative for the teacher in that they show her the limits of the child's cognitive ability, while for the constructivist (see below), errors demonstrate the child's way of thinking and so enable the teacher to provide teaching which takes the child on in her or his own way of thinking. Error analyses of written test scripts (Somerset, 1987; Hargreaves, 1995) could be used formatively from either of these perspectives.

### ***The transmission model, constructivist model and sociocultural model***

Brown (1989) contrasted a *transmission model* of learning, which assumes that the learner is mainly dependent on the teacher with a *constructivist model*, which assumes that knowledge is built up by the pupil in the form of constructed schemas. The *transmission model* is of the teacher owning knowledge and skills which she passes on to passive learners, perhaps many passive learners. As with the behaviourist model, knowledge and skills appear to lie outside the control of the learner until the teacher passes them on. This often leads to the practice of specific information being prescribed for various stages of the system, again implying that knowledge is quite independent of the learner. The teacher's job, therefore, is to make sure pupils take in the right information. The form of assessment which will enable teachers to evaluate how much information a learner has learnt, or memorised, could be a test in which the pupil is asked to reproduce certain facts or display mastery of certain skills. A good assessment will be one which faithfully reflects what has been taught.

Assessment of transmitted information may be of a normative type, comparing pupils against each other rather than against learning criteria. Harlen and James (1997)



pointed out that normative assessments may not be useful formative tools in terms of assessment for learning. Whilst norm-referenced assessment might help teachers recognise the existence of a problem, it can offer them no help in knowing what to do about it and ‘... may simply have a deleterious effect by labelling or pigeon-holing pupils’ (p.371). In addition, Ames (1992) showed that children felt more negative about themselves when

... the classroom climate is focused on winning, out-performing one another, or surpassing some normative standard, than when children focus on trying hard, improving their performance or just participating (p.264).

In the constructivist model, of which Piaget (1934) is sometimes considered a leading proponent, learning is a process whereby students take in information, interpret it and connect it to what they already know. Learners ‘... construct and then reconstruct mental models that organise ideas and their interrelation’ (Shepard, 1991, p.8). Constructivist views of learning are based on an interpretivist epistemology. As with interpretive approaches, the constructivist paradigm does not accept that reality is fixed and independent of the observer, but that there are multiple constructions of reality. Each individual, then, constructs knowledge independently. Social constructivists emphasise the socially situated nature of much learning (Pollard, 1990).

The teacher who holds constructivist beliefs tries to find out how each child is currently perceiving reality and making meaning out of it. This demands an individual approach to learning. Having found out how the child’s thoughts connect to each other and to the environment, she starts from the child’s viewpoint and helps correct or deepen the child’s thought connections. In assessment for learning, seen from the constructivist perspective, the assessor’s role is to make an immediate decision about how best to take the child on. Since the constructivist model of learning is based on the notion of relationships between ideas in the child’s mind, the context in which learning occurs is an essential feature of learning and its assessment (Gipps, 1996). Formative information will concern not only skills mastered, but how the child approached the task, how frequently she or he could repeat performance, in which contexts and with how much and what sort of support: the focus is not only on learning outcomes, but on the processes of learning itself for the enlightenment they give about the child’s constructions of knowledge.

*Authentic* assessment, popular in the United States in the 1990s (Maeroff, 1991), is based on the notion of the importance of context and process. McLean (1990) defined an authentic test as one which faithfully reflected the conditions under which achievement normally took place; and the knowledge and the processes essential to the subject matter being assessed. Because of the importance of context, authentic assessments may include a portfolio of evidence for a child's typical, best or self selected achievement or progress, taken from everyday classroom activity; as well as *ephemeral* evidence in the form of teacher notes from observation and discussion (Sizmur *et al*, 1995). Again, each child may produce different outcomes from another and under different circumstances, therefore a variety of assessment techniques will be necessary for effective assessment (Brown, 1997).

An important point, made by Kimberley *et al* (1989), Porter (1989), Sadler (1989) and others was that norm referenced assessment could motivate those at or above the norm; but a system tailored to the individual context could also motivate children below the norm. The same was said for criterion referenced or graded assessments: even if children work at their own pace, some will fall far behind and so become demotivated. By individually negotiated assessment, they could better *own* the goals of learning.

Broadfoot (1979) suggested that assessment would contribute to learning if '... undertaken by both teacher *and* pupil... [since it] may help to reinforce the awareness of both teacher and pupil that learning is essentially an interactive activity' (p.7). That is, the teacher has to learn how the child is thinking in order to teach. In this sense, the lop-sided power relationship inherent in learning through transmission, is improved. Teacher depends on pupil as well as the reverse. Black and Wiliam (1998) likewise suggested that assessment could only be effective formatively if the pupil her or himself were involved in the process of assessment and feedback. Self-assessment, however, is only likely to be effective or possible where it is genuinely valued for its own sake, not only as a means to trying to make the child *feel* more in control (Sadler, 1998).

The difficulties of assessing how children are constructing their own meaning, compared to assessing how well they have remembered information from outside, have been noted (Brown, 1989). Like any individual assessment, the use of portfolio assessment is time-consuming and 'Pupils may be unwilling to engage in the intellectual

process, while the teachers may be unable to give up control of the discussion' (Gipps, forthcoming). Torrance and Pryor (1998) pointed out some of the difficulties in questioning young children, in order to find out their perceptions. Questioning as a means of assessment is one way in which the child learns that the teacher has the power in the classroom. Rogoff (1990) described how

In some cultural settings, however, the appropriate behaviour may be to show respect to the questioner or to avoid being made a fool of by giving the obvious answer to a question (p.59).

Socioculturalists, on the other hand, see learning as dependent not only on the teacher or on the individual; while assuming human agency in the process of constructing knowledge as do constructivists, they argue that meaning derived from interactions is not exclusively a product of the person acting (Bruner and Haste, 1987). They see learning not as the acquisition of knowledge but as increasing access to participating roles in expert performances so that learning is not a way of coming to know about the social world, but a way of being in it. From such a perspective, assessment has to be dynamic, with the assessor playing a role within the assessment itself. Vygotsky (1978) suggested that assessment was appropriate for formative purposes if it allowed the use of auxiliary tools, including adults and so produced best rather than typical performance. Such an approach to assessment is antithetical to the traditional psychometric model of testing in which any collaboration during testing is considered to be cheating. The dynamic sociocultural model carries with it messages that the pupil may negotiate with the teacher and so, in an environment where this is culturally permissible, suggests a further shift in the power relationship between teacher and learner. Broadfoot and Osborn (1987), however, pointed out that where the power relationship is negotiated between teacher and state, and teacher and pupil, everyday life in the classroom could become increasingly stressful. They asked,

Who had the greatest crisis in morale - the French *instituteur* who, tied to a traditional curriculum which may have little relevance or appeal for many pupils, has neither the right nor the skill to change it, or the English primary teacher who sets herself the unachievable goal of an individualised pedagogy in a class of over 30 pupils and who had no clear idea of the limits of her responsibility? (p.300).

### *Surface-level and deep-level learning*

If learning is perceived to be imported from outside, its impact could be expected to be less deep than if it is perceived to happen from within the learner. The behaviourist and transmission models of learning could be expected to promote surface-level learning. The cognitivist and constructivist models could be expected to promote deep-level learning.

Crooks (1988) suggested that deep learning embodied an intention to develop personal understanding while surface learning reflected the intention to be able to reproduce content as required. Deep learning involved active interaction with the content, particularly in relating new ideas to previous knowledge and experience while surface learning expected a passive acceptance of ideas and information. With deep learning, ideas were linked together using integrating principles but with surface learning, there was a lack of recognition of guiding principles or patterns.

With deep learning, children related evidence to conclusions while with surface learning, they focused on assessment requirements. When something is learnt with deep learning, it is actively understood and internalised by the learner and makes sense in terms of a learner's experience of the world, rather than being a collection of isolated facts which have been memorised. Deep learning differs from rote learning essentially in that it is linked to previous experience and so can be used in situations different from that in which it was learnt (Harlen and James, 1997).

Assessment which is formative for surface-level learning might involve the teacher in constantly checking that information had been memorised, using closed oral questioning, asking pupils to recite, written examinations including questions demanding short answers and, perhaps, multiple choice and matching questions. Harlen and James (1997) described the distinctive features of assessment founded on a belief in deep-level learning:

... concern for deeper learning and the constructivist view of learning have done much to draw attention to aspects of pupils' thinking that teachers should take into account. Gaining access to the ideas and mental frameworks that pupils have already formed, accurate or otherwise, is an integral part of teaching for understanding and requires teachers to adopt new strategies for lesson planning and different forms of questioning (p.371).

### ***Basic skills and facts and higher order skills***

Bloom's (1956) *Taxonomy of educational objectives* depicted learning as becoming increasingly more complex and abstract, starting with simple knowledge acquisition. Inherent in his view was belief in a hierarchy of learning, starting with learning dependent on few mental connections and moving up to those involving complex connections. As with the behaviourist model of learning, from such a perspective, formative assessors find out what level of learning the child is at, in order to give instruction at the next level up. Testing initially assesses basic facts and skills and increasingly looks for more complex cognitive connections. In this model, the focus is specifically on how children learn, although some theories about learning dismiss *higher order skills* completely within the learning paradigm. That is, the learning of *basic skills and facts* is the goal of education.

*Basic skills learning* implies that mastery of some elementary skills are essential for more advanced learning, and also, possibly, for life. In other words, *basic skills* are the lowest skills on a hierarchy. Bracey (1983) suggested that

People who speak of basic skills usually believe, either implicitly or explicitly, in the existence of an agreed-upon set of skills that are fundamental to other skills and essential to survival in today's world. (p.718).

The notion behind assessment of the basics is that, as both population and compulsory primary education increase, a minimum learning level of basic skills should be guaranteed (Porter, 1989). Assessment of the 'basics' usually refers to testing 'basic' skills in the first language and mathematics, although it is noteworthy that the UNESCO-UNICEF international monitoring programme added *life skills* (Chinapah, 1996); and in Korea, the Korean Educational Development Institute considered the three basic areas to be: verbal ability, computational ability and utilising material and data (Chan-Jon Kim, 1996). These examples show that learning of *basic skills* may involve assessment of complex cognitive connections as essential in some aspects of language and mathematical learning. The approach, though, implies a belief that initially students make fewer cognitive connections than later on in the learning process, rather than implying that initially, connections are made differently than later and less appropriately.

Minimum competency testing in the United States of America has become identified with assessment of the basics and was implemented in response to perceived lowering of standards there (Madaus, 1985; Torrance, 1984). It revolved around state administered testing of a hierarchy of basic skills. With this approach, having assessed the child's mastery of a lower set of the basic skills, the teacher makes the formative decision about which task in the hierarchy the child is ready to attempt next in order eventually that she or he become *minimally competent*. In this way, teacher and pupil adhere to a clearly defined structure and no child should be able to pass through school without having learnt important 'basics'.

One criticism of assessing the *basics* is that these vary across time, culture and individuals (Berlak, 1992; Bracey, 1983; Murphy and Torrance, 1988). Critics of minimum competency testing stress that 'Those same minimum standards that work for some may stifle others' (Porter, 1989, p.349) and also that standards are based around the '... lowest common denominator' (Airasian and Madaus, 1983, p.105). Somerset (1982) quoted a Kenyan teacher saying of his pupils,

Many teachers know that their students should be able to do more... but their main concern is to get the students through the examination - and we can see that the teacher is judged on the examination performance of his students (p.18).

Cole (1990) argued that the basic skills model was not inappropriate but insufficient: the *higher order skills* model needs to be integrated with it and appropriate language used to indicate that integration. Higher order skills are described as the ability to '... implement critical and creative thinking, discover solutions and participate in debate using analytic and deductive reasoning' (MoE, 1992, p.53). Land (1997) described how a good assessment, when higher order skills were valued as primary targets of learning, focused on '... core types of learning that recur across the curriculum: conceptual understanding, knowledge representation, problem solving, communication, and team work' (p.6). Glaser and Silver (1994) suggested that in assessment of higher order skills and within a sociocultural paradigm, the pupil observes how others reason, and can then receive feedback on her or his own efforts. Formative assessment, from this perspective, focuses not on content, but specifically on the processes by which the pupil arrives at solutions, believing that once the tools for learning are internalised, any content can be learnt deeply and so applied to a range of contexts. Thus, test questions

tend to ask 'why' and 'how', instead of 'who', 'where', 'when' and 'what' (Somerset, 1982).

#### **2.4.4 Assessment as motivation for learning**

However learning is perceived, assessment may provide motivation for learning. The question is whether motivation is an agent for the improvement of the individual's learning or whether it damages it; and what the broader social consequences are. The role of assessment in the child's perception influences the extent and effect of motivation. *High stakes* selection examinations may encourage *extrinsic motivation*, while *low stakes* assessment for learning may encourage task interest to a greater extent. If, however, assessment for selection acting as an extrinsic 'carrot' (McNamara, 1982) later encourages intrinsic task interest as Allport (1965) suggested, does assessment within learning overall inspire less motivation to learn than assessment for selection? Where external examinations were removed in Mexico and in China, for example (Little, 1984), demoralisation set in. In this case, then, reform of examinations and the curricula they assess will inspire more 'learning' than 'assessment *for* learning' (perhaps in the form of certification) in which the child perceives no external drive.

Broadfoot (1998) however claimed that '... the emphasis on measures such as national tests is, arguably, actively undermining the potential of many children to succeed. This is because confidence is crucial to learning' (p.25). Others have used tests expressly to provide a sense of success for lower-achieving pupils (Close and Brown, 1987). McNamara (1982), however, questioned whether 'carrot driven' learning towards even the best developed assessments could endure once the carrot holder went away. McNamara commented that

... it is relatively easy to delude oneself into confusing compliant behaviour with the internalization of new attitudes and value systems. Once the lid comes off, as with... the death of Mao Tse-tung, some of those who were the most active conformists tell us that they had no choice but to lie with apparent enthusiasm (p.52).

McNamara applied this comment both to 'draconic' and 'soft cell' sanctions: in the present case, high stakes assessments for selection and low stakes assessments formative for learning.

Alderson and Wall (1993) described the *Yerkes-Dodson Law* which said that

... an increase in level of motivation is accompanied by an increase in learning, up to an optimal point. However, beyond that point an increase in motivation seems to have negative effects and performance declines (p.119)

A test, therefore, needs to be high, but not very high, stakes for the individual. But how high stakes the test is perceived to be depends on each child's need for success and how the child perceives success to be achieved. The lower the need for or possibility of success, the more effect the inherent interest of a test will have.

Where teachers clearly use the threat of high stakes assessments to discipline children, as described by Lewin (1980) in Malaysia, this implies that the children are not intrinsically interested in their work. The teacher's own lack of intrinsic motivation in a situation where assessment regulations may be dictated from an outside examination syllabus, and her own beliefs about learning, may encourage her use of this external carrot and so create a vicious circle which is unlikely to encourage task interest (Porter, 1989). Cohen (1990) illustrated how motivated the pupils and teachers became in the Eight Year Study in the United States, where 30 schools were given complete curriculum freedom; although this was not the case when the same happened in Mexico and China.

More influential on children's motivation, however, as Payne and Barker (1986) pointed out, are parents' educational expectations. Seginer (1983) found that this applied across age groups and across races. Sadler (1989) stressed that a learner may, however, choose to own or to reject some external goals; only those owned by the child play a significant part in the voluntary regulation of performance. That is, it did not depend on whether the assessment was a high stakes examination or low stakes classroom investigation: what mattered was that the child wanted to succeed in it. Then, Sadler argued, the most motivating goals were *hard goals* which were

... specific and clear rather than general or vague, harder and challenging rather than simple or easy, and closer to the upper limit of an individual's capacity... Hard goals act to focus attention, mobilize effort, and increase persistence at a task (p.129).

The Students Learning Orientations Group's (1987) research in Japan, Malaysia, England, India, Nigeria and Sri Lanka suggested that great care should, therefore, be taken in making any generalisations about the relationship among learning, assessment and motivation, since it is highly context dependent (Little, 1996b). There seems to be



no clear indication that assessment within learning, as opposed to assessment for selection, certification or accountability, is more intrinsically motivating.

## **2.5 Literature overview: the roles of assessment in primary education**

This chapter has examined three official, stated purposes for educational assessment and considered the type of improvement each one is concerned with. Assessment for certification at primary level was shown to be often combined with assessment for selection, and in those cases, the assessment instrument tended to be the formal written examination. Assessment for selection was shown to have purposes and also effects, beyond the selection operation itself: selection within education could also be seen to increase the country's productivity; to single out a ruling elite; or to provide a route for social mobility. For these reasons, the objectivity of the assessment instrument was seen to be of paramount importance. The negative effects of selection examinations on pupils, parents and teachers were suggested, but also the potential of examinations to bring about educational improvement.

Assessment for accountability was, similarly, seen to be interlinked with assessment for system improvement, and was often instrumentalised through monitoring exercises. While system improvement was depicted as one essential purpose for assessment, some negative effects of such exercises were shown with particular reference to teachers. Classroom practice and morale could suffer when the emphasis in the system was on accountability rather than, for example, the encouragement of local initiative.

It was suggested that assessment for the improvement of the individual's learning was less likely to be an official, stated purpose for assessment because formative purposes for assessment concern the individual pupil, teacher and classroom processes rather than official policy-makers, who would be more concerned with the outcomes of assessments which could be reported in a summative way. Formative purposes were those of constantly improving the teaching and learning processes in the classroom and therefore depended on how those processes were perceived. Different assessment instruments would be needed if learning was believed to be bestowed on the learner from outside, than if it was believed to be an internal process of making sense of the outside world.

Finally, motivation was shown to be important in assessment, both as a purpose and as an effect. The relationship between assessment and motivation was, however, shown to be complex and depended on such factors as whether the motivation was extrinsic or intrinsic, how *high-stakes* the assessment was, how near the *norm* the pupil felt and parents' expectations.

## **Chapter 3. Research methods**

### **3.1 Introduction**

This chapter describes the methods of research which were used in the current research into the roles of assessment in the Egyptian primary school. To start the chapter, there is an outline of the research questions and of the methods chosen to find answers for each one. Then the main elements of the field research are described, including the research diary; how the case study approach and also school were chosen; the pilot study; the writing tasks undertaken by pupils; interviews; observation; finally, validation. The chapter ends with some discussion of practical issues related to the field research. Throughout the chapter, extracts from the research diary are provided in order to make this account of methods vivid.

### **3.2 Research questions and methods: rationale**

The main research question addressed by this thesis is, what are the roles of educational assessment in primary education? In order to find answers to this overriding research question, the case study approach was chosen (see below) and field research was carried out in Egypt, mainly in a primary school called El Nasr Girls College Junior School in Alexandria. Four subsidiary research questions guided the field research. These were more focused than the main research question, and related specifically to the case study primary school. These subsidiary research questions are displayed in Chapter 1.2 and the research methods used in answering each one are described below.

#### **3.2.1 Appropriate methods of research**

Distinctions between qualitative and quantitative research paradigms have been made (Lincoln and Guba, 1985). In terms of epistemology, the qualitative paradigm has been described as naturalist or interpretivist, where realities are believed to be multiple, constructed and holistic; while the quantitative paradigm is positivist and reality is believed to be single, tangible and fragmentable. Lincoln and Guba (1985) also described the following distinctions.

- **For the qualitative researcher**, knower and known are interactive and inseparable; only time and context-bound working hypotheses (idiographic statements) are possible; all entities are in a state of mutual, simultaneous shaping, so that it is impossible to distinguish causes from effects; inquiry is value-bound.
- **For the quantitative researcher**, knower and known are independent; time and context-free generalisations (nomothetic statements) are possible; there are real causes, temporally precedent to or simultaneous with their effects; inquiry is value-free.

The polarities which represent the two paradigms have been described in terms of words versus numbers, natural versus artificial settings, a focus on meanings rather than behaviour, adoption or rejection of natural science as a research model, an inductive versus deductive approach, identifying cultural patterns versus seeking scientific laws and idealism versus realism (Hammersley, 1992). The distinctions have been criticised, however, either because they are less rigid than they are made out to be, or because one or other is considered untenable as a starting point for educational research. Collins (1992) considered two paradigms as too few and described a third, the critical paradigm; for example, feminist, Marxist or Freirean. She wrote,

... the goal of the positivist researcher is to predict; the goal of the interpretive researcher is to understand; the goal of the critical researcher is to empower and liberate (p.183).

Collins suggested that the research paradigm and therefore its methods depended on the goals of the research. The goals of the current research were to predict certain effects and make some generalisations, believing that inquiry may have value-free elements. For example, the current research sought to predict some effects that certain assessment policies might bring about and to generalise this relationship to assessment policy and practice in countries other than Egypt. However, its further goals were to understand how individuals experienced and used assessment in their daily lives and to understand how they perceived central reforms. Hopkin (1992) described this focus on ‘... understanding the culture and context of the research in its natural setting’ (p.134) as an important indicator of the qualitative nature of the research. Because of this goal of understanding experiences and perceptions within the culture and context of the natural setting, the case study approach was chosen for the current research. As for empowering and liberating, no such ambitions were attached to this research, although

as Chapter 1 explains, the data it exposes might empower the making of appropriate decisions which take full account of assessment's potential uses and positive and negative effects.

In these senses, the current research does not easily fit into either the qualitative or the quantitative category if such a distinction has to be made, although to use Hammersley's opposites, its emphasis was on words and meanings more than numbers and behaviour, natural more often than artificial settings, an inductive more than a deductive approach and the identification of cultural patterns rather than scientific laws; which suggests that its paradigm was more interpretivist than positivist. In some respects, it may be said to belong to the phenomenological approach within the hermeneutic, qualitative tradition, in that priority was given to *actors'* accounts of social reality (Schutz, 1967). For this reason, interviews and pupils' own writing were principle methods of inquiry. The analysis of documents as a research method could be seen as a more scientific collection of tangible data, and yet their meaning to pupils, parents and teachers was one goal for collecting them in addition to writing a separate account about the data. These research methods were therefore chosen as most appropriate to achieve the goals of this research:

- Analysis of writing
- Analysis of documentation
- Interview
- Observation.

### **3.2.2 Research methods in relation to research questions**

These methods were appropriate in finding answers to the subsidiary research questions as follows.

#### **Subsidiary research question 1: methods**

*What were assessment's official roles in Egypt generally and in the case study school specifically?*

There were two parts to this question, one addressing Egyptian policy generally and the other addressing this policy as manifested in the case study school. To find answers to

each part, there were two methods of research: analysis of documentation; and interview. Since this question was about policy, answers were more factual and less connected to subjective perceptions than the other three questions.

- **Analysis of documentation**

Nationally prescribed syllabuses, text-books, *taqwim* (student self-evaluation guides), examination timetables and policy documents were analysed to find out what Egyptian schools were being asked by the government to do in terms of assessment. In addition, school documents were analysed, including teachers' termly and weekly study plans, test and examination result records, revision sheets prepared in school, staff appraisal guidelines and reports and school arrangements for examinations. The aim of analysing these documents was to find out how policies were being practised in the case study school and what other practices were prescribed by the school management to assess children and teachers efficiently. The information gained from this method informed Chapter 4 and Chapter 5.

- **Interview**

Personnel in positions of authority with regard to assessment, were interviewed, from the MoE and the National Centre for Examinations and Educational Evaluation (NCEEE) in Cairo and the Centre for Curriculum and Instructional Materials Development, the local inspectorate in Alexandria and the management within the school itself. These people were asked to describe in detail the official national or school policy on assessment and how it was practised. The information gained from these interviews informed Chapter 4 and Chapter 5.

### **Subsidiary research questions 2 and 3: methods**

*In what senses did pupils, parents and teachers experience the assessment system positively?*

*In what senses did pupils, parents and teachers experience the assessment system negatively?*

Since responses to these two questions included the feelings and impressions of pupils, parents and teachers, analysis of official documentation was not an appropriate research method in finding answers to them.

- **Analysis of writing**

Whole classes of pupils were asked to complete some sentences relating to their experiences of examinations so that they could express themselves directly. They were also invited to write about the way they spent an average weekday 24 hours, so that the practical details of their daily experiences around school could be analysed. Some essays from an English language examination were also read and analysed for what they showed about pupils' feelings about examinations, but this time, pupils did not know their writing would be analysed in this way. Information gained from each of these sources informed Chapter 6.

- **Interview**

Only pupils had their writing analysed, but pupils, parents and teachers were all interviewed. Through individual interviews, or interviews with mother and father, it was hoped to learn directly about each party's experiences of assessment in the school and national system. It was also hoped to reach greater understanding of what they were saying by visiting some of their home surroundings. By interviewing the three parties separately, their respective responses could be compared and contrasted. Responses from these formed the basis for Chapters 6, 7 and 8.

- **Observation**

Pupils and teachers were observed during lessons and examinations and teachers were observed during their marking of examinations. This was so that the experiences pupils and teachers described in writing and in interview could be triangulated and a more comprehensive picture of the system could be developed. Information from observation mainly informed Chapter 5.

#### **Subsidiary research question 4: methods**

*What were pupils', parents' and teachers' experiences and perceptions of central reforms to assessment?*

Pupils', parents' and teachers' experiences and perceptions of central reforms to assessment were sought using all four research methods. Policies in national and school documents and as described by people in authoritative positions, were considered first. Pupils' writing was reviewed in relation to these policies. Interview responses and

evidence from observations were then analysed in the light of these policies. In interview, parents and teachers were asked directly about their experiences and perceptions of central reforms to assessment. Information from these sources was used for Chapters 4 to 8.

### **3.2.3 Main elements of research methods used**

The following were the main elements of how these methods were employed in the field research, once the research questions and research approach had been decided.

- **Research diary: January 1996 to June 1996 (six months)**
- **Finding the case study school: February and March 1996 (two months)**
- **Pilot study: March 1996 (two days)**
- **Analysis of writing: April and May 1996 (two months)**
- **Interviews: April and May 1996 (two months)**
- **Observation: December 1996 (one month)**
- **Validation: December 1996 and June 1998 (two months).**

### **3.3 Research diary**

The diary was an outlet in which to note down issues which seemed important both in the specific context and more generally. It allowed for constant reflection or *iteration* (Vulliamy *et al*, 1990), on data, methods and on feelings about both. This reflection related to issues of research and being a researcher in a less-industrialised country, specifically Egypt, in addition to the assessment issues which the research addressed directly. The research diary was a record of progress in terms of the practical research and theoretical thinking, which helped in the drawing out of conclusions and implications. It was also one of the main sources of data for this chapter.



### **3.4 Case study approach: the case study school**

#### **3.4.1 The appropriateness of the case study approach**

##### *Goals of the research*

The goals of the research are re-presented above. In summary, the field research was to determine how pupils, parents and teachers used and experienced the assessment system in their school; that is, how they behaved in relation to it, and how they perceived and felt about it. The aim of the research was also to examine how they accommodated official policy, in practical, theoretical and emotional terms.

##### *Uses of the case study approach*

The case study researcher typically *observes* the characteristics of an individual unit - a child, a clique, a class, a school or a community. The purpose of such observation is to probe deeply and to analyse intensively the multifarious phenomena that constitute the life cycle of the unit with a view to establishing generalisations about the wider population to which that unit belongs (Cohen and Manion, 1985, p.120)

The case study's strength lies in its attention to the subtlety and complexity of the case in its own right including the discrepancies or conflicts between the viewpoints held by participants. The aim of a case study is to '... illustrate relationships, micropolitical issues and patterns of influences in a particular context' (Bell, 1993, p.9). The way in which the study is reported is important, since it must not only be a realistic picture of how the case study unit was, but should be vivid enough also to allow for the reader to make a personal interpretation. Representation rather than evaluation is its purpose. Vulliamy warned that

...critics of case-study evaluation studies have suggested that a more general application of their findings is frequently limited by their explicit emphasis on evaluation rather than on understanding (p.228).

A case study may be representative of other like units or may be worth studying because it is unique. Statements emerging from case study research may, therefore, be nomothetic or idiographic. In the current study, generalisations or rules relating to the processes of assessment were sought and hypotheses about these tested. Yet, Egypt had unique as well as shared aspects in its assessment system and little research of any kind has been done in Egyptian primary schools. Therefore, the exploration could be open and new theories generated.

For the current research, the management of the case study school needed to allow the research goals to be reached and therefore the chosen research methods to be used. This case study had to be, first of all, an Egyptian primary school. The management had to allow the researcher enough flexibility and access to come to an understanding of the culture of the place in its natural setting so as to generate theories as well as test them out. The research methods' emphasis on words meant that language had to be accessible and people needed to be willing and available to express themselves. There needed to be opportunities for interviews with pupils, parents and teachers and for observation of some lessons. Children's writing and school documentation had to be available.

### **3.4.2 The appropriateness of the case study school**

The original aim was to work with the NCEEE in a government primary school in Cairo, but access was still being denied by the end of February and time was short, since the examination period started at the beginning of May. The research diary for February 4th, 1996 referred to DC, who

... more or less founded the NCEEE in Cairo. He told me some useful general facts about education in Egypt, and the upshot is that I am going to Edinburgh on Wednesday to pick up the documents he has about the whole project. He said that access into schools is notoriously difficult, especially because teachers and governors are not happy with the quality of education they are providing... He gave out warning signals about getting access to anything there, and about people sticking to their promises.

For this reason, the headmistress of El Nasr Girls College in Alexandria was asked to host this research and she agreed immediately. The headmistress of both the whole school [ID] and the Junior school [QM] were acquaintances of my family from the 1950s and I myself knew them from my work at EGC in the 1980s. The requirements were discussed initially by telephone and fax, and then further on my arrival in Alexandria in March 1996. Following the decision to use EGC as the case study school, the research diary read,

[February 19th] ... the whole project will be much more practicable: I won't have to translate everything, find schools, find ways of getting people to talk...

El Nasr Girls College was an Egyptian primary school. Unlike state schools, though, it was owned by a syndicate, including some government officials, rather than being

totally state-owned. However, its curriculum and assessment arrangements were the same as for all primary schools in Egypt; only its intake was above average in terms of social privilege since some fees were charged and some English was used.

Because I had past connections with the school, I was given flexibility and access in my research. The headmistress welcomed me particularly because she tended to

... revere western methods because they are western. Her own experience in England was in private schools with loads of money and few children - very different from EGC. But she is interested in and receptive to ideas that may make education here 'better' [Research diary, March 19th, 1996].

My own previous experiences in Egypt also helped me come to an understanding of the culture of the place in its natural setting.

The ECG was different from many Egyptian schools in that half the pupils' schooling was through the medium of English. While this made it atypical of Egyptian schools, its assessment system and curriculum materials were the same as for all other Egyptian schools, albeit some materials were translated. But the English medium did facilitate my own deeper understanding of what people were saying. Although my command of Arabic was quite good and became better, having the means to communicate in either or both of the languages, English and Arabic, enabled me to understand meaning more fully. In the English medium school, I could always find someone to translate an Arabic phrase I wanted to understand. Sometimes, during translation, I received elaboration about broader meaning, as well as word for word translation.

One problem with the bi-lingual nature of the school was that personnel in authority encouraged me to talk with pupils, parents and teachers who had the most proficient English. I had to ensure that I talked to a mixture of fluent English speakers and non-English speakers. On March 21st, the research diary read:

Parent interviews do not seem to be a problem, except that I must get 'random' parents: they were talking about selecting co-operative and English speaking parents, but this must be avoided. I told them about tracking six children in each class, and talking to their parents, and this seemed fine, so I must make sure these six children are randomly selected from a class list. But the school is happy to provide a venue for parents to come in and be interviewed...

ID and QM agreed that children could be taken out of class for interviewing and that teachers could be approached. They decided on appropriate classes for observation,

whose teachers would be willing to give support. It later transpired, though, that the pupils and teachers concerned were never approached by the headteachers and they simply assumed that, as a foreign researcher, I had the headteachers' permission to intervene. The attitude is reflected by this account of teacher SMY:

SMY says they are all afraid of me. I heard her say to REB, a young teacher I interviewed yesterday: 'Why are you afraid of foreigners? They are just as simple as we are; perhaps more so' [Research diary, May 9th, 1996].

### **3.5 Pilot study**

The pilot of the single case study is difficult, in that no other case may be like the case to be studied. In some senses, the pilot is less crucial than for the pilot of a test which is to be nationally standardised, for example, since the flexibility of case study gives the researcher some opportunity to learn during the process. For the current research, a pilot study was carried out in a school which had similarities to the case study school; but learning during the process was also an important means to explore the potential of the research methods and instruments.

The following extract from the research diary describes the main practical as well as methodological aspects of the pilot study. The methodological aspects were testing the coverage, administration and timing of various interview and observation schedules, and the practical aspects were getting a feel for the accessibility of informants. The study was carried out in Kaumaya primary school which was a private, English medium school whose headmistress had once been on the staff of EGC. On March 25th, the research diary read:

I went to Kaumaya today, to pilot some interview and observation schedules. I did interview B with DT, which took at least an hour, much longer than I'd estimated, and I found that interview C was partially covered by default... I then went to see the head, IB, who used to be my head of department at EGC in 1986, and she invited me back to interview her in 15 minutes. When I went back, there were streams of people in her office, to whom she gave attention to the last-arrived, a very Egyptian way of ordering. A father had come to complain that his son had been standing outside the classroom for two weeks... So much was going on in her office, however, that I took leave and was invited back any day 'early in the morning'...

I then tried out interview D with two mathematics teachers in Junior 1 and 2... My schedule worked all right, except that again it took for ever, and also, they were ready to chat and to give me unsolicited details, which were interesting but not strictly pertinent to the schedule. It made me realise how long all my interviewing is going to take, and how much there is to find out; and also how I must be focused...

At the end of the interview, I attended one of the teachers' mathematics classes to try out my observation schedule... Again, it made me realise how much goes on in a class: it is essential to focus. Before I get to real observation, that schedule is going to need a lot of revision... All sorts were going on in the class today: a peculiar mix of formality and intimacy...

The observation schedule referred to in this extract never proved satisfactory. In this sense, the observations made in April and May, 1996, were pilots for the intense period of observations made in December, 1996, by which time the observation focus and schedule were fully prepared. The interviews improved with practice, but the few piloting attempts prior to their actual administration in school, proved enough to make even the first interviewees' responses at EGC valuable.

The pupil interview was piloted on the daughter of an acquaintance from outside school and the parent interview was first piloted on her parents. The parent schedule then had to be translated and practised with Jaida, the Arabic tutor I took lessons with. She was an eighteen year old pupil at the EGC and so could provide particularly useful practice for me, in asking interview questions and understanding responses to them.

### **3.6 Analysis of writing**

#### **3.6.1 Aims for writing tasks**

The aim of asking children to write was twofold. Firstly, the intention was to find out their experiences of examinations and the practical details of their daily experiences

around school. Secondly, writing was used rather than just interviews, so that a broader, more quantitative, understanding could be gained about a larger number of pupils' experiences in a relatively short space of time.

There were three writing exercises: whole classes of pupils were asked to complete some sentences; one whole class were also asked to write about the way they spent an average weekday 24 hours; and some essays from an English language examination were read and analysed, although pupils did not know this writing would be analysed for the research. Responses from the three exercises informed Chapter 6.

### **3.6.2 Sentences to complete**

Each child in two classes received a photocopied sheet containing sentences to complete. Fewer and simpler sentences were given to grade 3 pupils than to grade 4 pupils because of grade 3 children's more limited language skills. The children were told to be completely honest; their class teacher explained that the researcher wanted to know what they really thought. The teacher wrote on the board any words the children wanted to know. It was stressed that pupils could leave out questions they did not understand.

The teacher's control of the exercise may have caused some children to write to please the teacher. The children's familiarity with test papers where answers are right or wrong may have encouraged this. The teacher's familiar style of presentation and ability to respond appropriately to questions was however considered more important than keeping the answers confidential from the teacher.

When analysing responses, where a pupil had obviously misunderstood a sentence, her response was either ignored, or relocated to another sentence if appropriate. As with the analysis of all interview responses made in English, an interpretation of non standard English had to be made when analysing the written responses, but this was done with a knowledge of the Arabic language and the way in which Egyptian people use English.

### 3.6.3 Time schedules to fill in

On March 27th, the research diary read,

Reading Hilsum and Cane's (1971) 'The teacher's day', an idea came to me to give each of the interviewed parents a time schedule to fill out, as to what their child is doing for each hour of the day. The hours in the day would be listed, and on each hour, the parent would fill in the child's activity. Would it work?

Later the same day it read,

I have thought through and changed my mind about the parent time schedules. It would probably be more effective to get the pupils themselves to fill them in. I would start with one child, perhaps the one I sit with all day at school, coach them in how to fill in the time schedule, then discuss it with them on return.

Even this revised idea proved too difficult so that in the end, one whole class of children was asked to fill in time schedules while they were sitting in class, in retrospect to the past 24 hours, rather than recording as events actually happened. I piloted the approach with the five grade 3 children who were also the grade 3 interviewees.

A photocopied sheet was distributed to the same grade 4 children who completed the sentences about examinations. The sheet listed each hour of the day, starting with 6am, and a space was given next to each time. The girls were asked to write what they were doing at each hour of the day in the 24 hours up to 6am of that day, which was in April 1996. They were told to approximate to the hour, or to fill in the half hour if they wished.

### 3.6.4 Examination essays

All grade 4 children were asked to write an examination composition entitled *summer holiday* during their May English language examination. While marking these scripts with the grade 4 teachers, it was possible for me to note down some enlightening examples. In addition, the school gave permission, under tight security, for the grade 4 interview pupils' scripts to be identified and copied down.

## **3.7 Interviews**

### **3.7.1 Overview of the interview samples**

Interviews were held with the following people:

- **Six pupils from each of grades 2, 3 and 4**
- **Twenty of the same pupils' parents (bar one pupil)**
- **Nineteen teachers of pupils in grades 2, 3 and 4**
- **The headteacher of the whole school and of the primary school**
- **The school's mathematics inspector, who also wrote the external examinations**
- **The USAid trainer for the Centre for Curriculum and Instructional Materials Development**
- **The First Undersecretary for General Education at the MoE**
- **Assessment personnel at NCEEE.**

### **3.7.2 Aims for the interviews**

Interviews with all these parties were held in order to find out what the assessment policy was in Egypt and how the assessment system worked at primary school level. Interviews with the first three parties also had the aim of finding out how pupils, parents and teachers used and experienced the assessment system in their school; that is, how they behaved in relation to it, and how they perceived and felt about it.

Numbers of interviewees had to be limited to allow time for the parents of the selected pupils, as well as pupils themselves, to be interviewed before the examinations began at the end of April. Teachers had to be interviewed before term finished at the end of May.

### **3.7.3 Pupil interviews**

#### ***The pupil interview process***

Pupil interviews were carried out during April 1996, during the school day. They were held with the individual in a place where other people could not listen, such as an empty office, a bench in the garden or a deserted communal sitting area. Interviews lasted between twenty and forty minutes, depending on how much the child said. Responses to these interviews mainly informed Chapter 6.



The schedule is shown in Appendix A. The interview schedule was used as a basis for questioning, but pupils were encouraged to give any extra details they wished and not all questions were posed to all candidates. The research diary for March 28th read:

I have found it easier to use my interview schedules as prompts, rather than filling in spaces on the schedule, and rather than adhering to any strict order. I find that if something interesting has come up, then it's good to let it flow, and in the process, other prompts may be covered.

The child's comments were written down and typed up later the same day. If the child seemed to be having difficulty understanding or speaking in English, Arabic was used by the researcher or the pupil, or both, as appropriate. The research diary of March 29th read:

First two lessons, I interviewed two of my Junior 3 pupils, these two chosen especially because they are weak in maths. [Teacher] SEY had said it would be difficult as they probably wouldn't speak good English if they were poor in maths. [Pupil] Ahlam got more and more lyrical, and often in English, as the interview went on. And [pupil] Rania spoke not a word of English, but told me all about how it feels to fail and to work over the summer, and to think you've failed the repeat exam too.

### *The pupil interview sample*

Table 1 in Appendix B shows the name [invented], grade and age at time of interview of each child, as well as fathers' and mothers' occupations and each pupil's siblings. The table shows that the children's ages ranged from seven years and seven months to 11 years and three months. They all had at least one sibling and their parents tended to be professionals, including several of the mothers. Table 2 in Appendix B gives some indications about each child's personal qualities: whether they were vocal in interview; how they performed in the end of year examinations and unprompted comments from their English and mathematics teachers. The table shows the variety of qualities and abilities among the 18 interviewees. The comments given by teachers add some perceptive details about them, often including the role parents played in each child's school life.

### 3.7.4 Parent interviews

#### *The parent interview process*

The aim was to interview at least one parent of each of the 18 interview pupils. The headmistresses at EGC were ready to help organise parent interviews. The head of the primary school, QM and the head of English for the primary school, SES, vacated their offices when possible so that parent interviews could take place there. The grade 2 mathematics teacher, SMY, telephoned some of her pupils' parents and asked them to come in to school for interview. In grades 3 and 4, interview pupils were asked by their teacher in class to ask their parents when they could come for interview. There was an initial rush of parents who came in for interview by these means and half the parent interviews were complete by April 10th. The other half of the interview pupils' parents, however, seemed reluctant. Then pupil Rania asked if the interview could be at her house, saying her mother could not go out. After this, it transpired that those parents who were weak in English did not want to come into school. By April 18th, the research diary read:

I have another 'home' visit tomorrow with Ahlam's parents. It suits me well because then I get to see their homes and meet other members of the family. Today, for example, I was shown a wonderful Astroturf disco-roof where Samira can 'relax' after her exams. I think I am, after all, going to manage all the parent interviews I wanted.

Table 3 of Appendix B, showing the parent sample, shows that the four interviews conducted at home were in Arabic. Dina's mother and Nora's mother used Arabic even though they were in school. This was because Dina's mother came in very reluctantly after some persuasion by SMY; and Nora's mother worked in the school secretariat.

The research diary showed that from the researcher's viewpoint, 'parent interviews are stressful' [21st April]. This was partly because they lasted for between one hour and 90 minutes. I had to prepare the interview questions in Arabic, with Jaida my tutor. Then I taped the interview in case I missed meanings in their responses. Afterwards, I translated and transcribed each word with SMY or, sometimes, her 15 year old daughter, a long process which continued into the next research period of December 1996. On May 6th, 1996, the research diary read,

Another enjoyable day spent locked in the 'Christian' room [where non-Muslim pupils took their religious education], translating parent interviews with SMY. As yesterday, she interspersed her translation with her own comments and explanations which were as enlightening as the interview itself. She is genuinely interested in the subject matter...

The parent interview schedule, the responses to which mainly informed Chapter 7, is shown in Appendix A.

### *The parent interview sample*

Of the 18 interview pupils, 17 had at least one parent who agreed to be interviewed. The parent sample is shown in Table 3 of Appendix B. Magda, Ahlam and Samira's parents were interviewed as mother and father together. Amira's father was interviewed alone. Only Mira's mother refused to be interviewed, despite many invitations by Mira, myself and the grade 2 mathematics teacher. Mira's mother agreed to meet for interview but did not turn up. The school said she never came into school; but neither did she want to carry out an interview at home. It seemed that Mira's mother was fluent in English and confident about school matters since she was a teacher herself, but that she herself was very busy invigilating and marking examinations in the school where she worked. The situation put Mira in a difficult position so no more attempts were made to invite her, after May 22nd.

Whether interviews were at school or at home, if parents chose to let other siblings accompany them, this was acceptable. Where possible, though, the pupil herself was asked not to listen in to her parent interview so that parents did not feel constrained in this way, although with five pupils, this was not possible.

### **3.7.5 Teacher interviews**

#### *The teacher interview process*

By May, teachers only attended school part time, often arriving, signing a register, and leaving soon after. Once the examinations for grades 2, 3 and 4 were over, there were no children to teach and attendance was largely a matter of formality, dictated by the Ministry. This was a good opportunity for teacher interviews. Interviews were carried out individually in a place where other people could not listen, such as an empty office, a bench in the garden or a deserted communal sitting area. The teacher interview schedule shown in Appendix A was used as a basis for questioning, but teachers were

encouraged to give any extra details they wished and not all questions were posed to all candidates. Teachers' responses mainly informed Chapter 8.

Interviews lasted between fifty minutes and two and a half hours, depending on how much the teacher wanted to contribute. This meant that the maximum number of interviews possible in one morning was two. The research diary of May 8th read:

The grade 2 and 3 teachers sign in at 9, drink one coffee and go home. I tried to find teachers to interview after 10am and had difficulty. I did find two teachers and learned a lot from interviewing them, although the interview was over an hour long each.

Teachers' comments were written down and typed up later the same day. If the teacher preferred to speak in Arabic, as was the case with the three teachers of the Arabic subjects, Arabic was used for questioning and by the teacher. These Arabic interviews were tape recorded and translated with the aid of SMY's daughter later (since it was felt that SMY herself should not know what other staff had said). Not all teachers were keen to be interviewed. Some appeared nervous, many suspicious, some could not be bothered, but others seemed pleased to have the rare chance to give their opinions.

### *The teacher interview sample*

The teacher interview sample is displayed in Table 4 of Appendix B. A spread of experience was sought in terms of years of teaching experience in EGC, which ranged from less than one full year to 30 years. The 19 teachers were chosen from across the primary classes, but teachers of the 18 children who had been interviewed had priority. Of the 19 teachers, all except two had experience of teaching in at least one of the three grades, grade 2, 3 or 4. In addition to mathematics, English and Arabic, the subjects of science, physical education (PE), art and religion were represented to provide for a range of experience. It was noticeable that seven teachers had attended Al Manar school, originally run by Scots, when these teachers were themselves pupils, and only three were ex-EGC girls. Only three teachers had been educated entirely in Arabic schools; and three had attended the French medium school, Sacred Heart.

### 3.7.6 Interviews with headteachers

On arriving at EGC, I arranged weekly one-hour meetings with the headmistress, ID, and worked through the long interview schedule I had prepared, over the term. Its aim was to find out about assessment policy in EGC. My earliest experience with her was recorded in the research diary of March 21st:

We went over to ID's office, and she, by chance, was relieved of Mother's Day duties long enough to chat for a good half hour. She seemed lively, with smiling, youthful eyes: amazing for the head of nearly 5,000 children in a school which, she said, she didn't feel she could any longer keep the lid on.

The headmistress's time was short. These interviews made me realise how much my research relied on the headmistress's goodwill and the goodwill of all the other interviewees. The headmistress had nothing obvious to gain from having this research in her school. It seemed that she did hope to learn about the English system and at least keep some contact with it; also to promote the good name of the EGC and emphasise its specialness, to a western audience; and she requested that she could read whatever was written up on the basis of her interviews about the school, afterwards.

Because she was very busy, this meant not deviating from the schedule and looking more for factual answers than the headmistress's feelings, since she was the only person who knew some facts about how the school fitted into the Egyptian system.

Initially, I also interviewed QM, head of the junior school, about details of the system in the primary section of EGC. After a few days it became obvious that she was weary of answering my questions. The research diary of March 27th read:

... it sounds like it's going to be a constant struggle to pin people down. After all, why should they? What does QM get out of sitting for an hour or so telling me about the system? I shall have to work out how to... spread the attention around tomorrow, so as not to take up too much of one person's time.

After that, I tended to sit in the office of either QM or SES, head of English for the primary school, and absorb what went on. Staff, parents and pupils would come in to talk to them both. In the research diary of March 28th I observed that,

QM, so different from the English head who has absolute authority, and from whom both hostility and warmth will often be hidden, QM will argue, practically cry, laugh, sympathise and chastise almost all in the same scene.

I discovered that sitting quietly in their offices could be enlightening or at least provide a starting point for further questioning. I spent some time nearly every day, when I was waiting for interviewees, sitting quietly in QM's or SES's office. It seemed that SES, more than QM, enjoyed giving the unusual details about the system and took some pleasure in my witnessing some difficult confrontations in her office. Quite often we were able to laugh about them together afterwards and she would try to explain what was going on under the surface. Later, I shifted to sitting with SMY in the staffroom, since she had more time to spare and helped me arrange interviews with other members of staff at the same time.

One result of this 'sitting around' approach was that *my* thinking and the thinking of SES and SMY began to become confused. On 15th May I recognised:

I had a two and a half hour interview with SES, of whom I am becoming fond. She has ideas about what is wrong with the system which are close to mine and I'm not sure which way the cross-fertilisation has worked. I have spent so much time with her and she with me, our thinking is becoming interrelated. And the same with me and SMY [Research diary].

Information gained in this way, and in interview with the headmistresses, mainly informed Chapter 5.

### **3.7.7 Interview with the inspector**

The First Primary Mathematics Teacher, NMA, helped to arrange the interview with the first inspector for mathematics in Alexandria Central Zone who was EGC's primary school inspector. NMA and QM helped translate between myself and the inspector during interview, in case of difficulties in understanding. The four of us sat for about an hour in the dining room around a table, since the staffrooms had been evacuated for the external examinations.

The inspector gave details about his training, his work, how the system of inspectors worked, how the external examinations were written, how examinations differed across governorates and between English, French and Arabic medium schools, and recent or proposed national reforms to the assessment system. These data informed Chapter 5 as well as Chapter 4.

### 3.7.8 Interviews at NCEEE

For the most updated information about assessment developments in Egypt for Chapters 4 and 5, I interviewed the following people at the National Centre for Examinations and Educational Evaluation (NCEEE) in Cairo:

- Director of NCEEE
- Head of department for examinations development
- Head of department for research
- Head of department for training and dissemination
- Head of test development for mathematics
- Head of test development for English
- Test developer for English.

Access to the director was so difficult by telephone that in the end, I went to Cairo and turned up unannounced at the NCEEE. I had contacts in London who knew the director, but no contacts within Egypt. I described my journey to NCEEE on June 16th, 1996:

The Centre was miles and miles away on top of the Muqattam mountain. My taxi driver kept sighing, as each bystander he asked sent him further into Muqattam. Finally we arrived at a white block with soldiers at the gate. I went in and from then on the confusion and attention began. Because I had never been able to find [the director] in, every time I phoned from Alexandria, I had left messages only to say I was coming to Cairo this week. But no-one had the slightest clue who I was... They asked me to fill in my full details as a foreigner and it was comprehensively assumed I did not know a word of Arabic. [The director] was not in and no-one knew when he would next be in. They didn't or couldn't give me his home number, nor tell me any information before his arrival.

The staff were suspicious and eager to follow protocol exactly, when dealing with a foreigner. Luckily, the director telephoned the Centre just at that moment and I was able to arrange to meet him the next day. He arranged for me to meet the other heads of department. The rest of the time, I sat in people's offices, observing and asking questions. The staff there gave me information about how the assessment Centre was set up, what its function was, how it related to the Ministry and the Centre for

Curriculum and what reforms and changes were expected to the assessment system in Egypt.

### **3.7.9 Interview at the Ministry of Education**

Despite the efforts of influential people, it proved impossible to gain permission for an interview with the Minister of Education, Bahaa al Din. I wrote to him twice but received no acknowledgement. I telephoned his office frequently, which meant finding a telephone which could connect between Alexandria and Cairo since not all telephones communicated further than Alexandria; but I was always asked to try again later. In the end, the British Council in Cairo helped arrange an interview with the First Undersecretary for General Education, at the Ministry of Education. Information from this interview was intended to feed into Chapter 4 concerning national policy about assessment. Access had also proved impossible to the Ministry representative in the Alexandria governorate.

I felt very nervous walking into the Ministry in Cairo, but my overwhelming memory of the interview was of shock at the lack of formality of the interview and how the First Undersecretary responded to my research questions aggressively if at all. I sensed he had an antipathy towards the western education system I represented to him. The research diary of June 27th best describes his behaviour:



... at one point, he threw a whole file full of papers violently onto the floor in temper. He began our meeting, at 8.30 am, by telling me how unwelcome he had been made in England, how his tutor didn't offer him tea and his neighbours didn't want to talk to him... For a change, in his office when I spoke Arabic he was pleased and said, 'That's better, Arabic'. And yet, he was not interested in anything I had to say; indeed, he didn't really listen to my questions, let alone my comments. He spoke slowly and would not be interrupted, although much of what he said did not answer my questions. Twice I asked him what reforms were to be introduced soon and both times he talked about various irrelevant topics instead of answering...

He basically said he wanted no examinations in primary, that they were not necessary, that it was just teachers and parents who wanted them (*'What makes us fail is the parents and the teachers'*). When I pointed out that they were legally compulsory, he denied it. Shortly afterwards, he contradicted himself, and said that, yes, an examination at the end of the year was a good thing. ... This interview gave me little information as such; but it gave me a feel for the ... system at the top.

### **3.7.10 Interview with USAid curriculum development trainer**

A British expert in training for curriculum development, who was working for USAid at the Centre for Curriculum and Instructional Materials Development in Cairo, provided information about curriculum reform because of its close links with assessment practice. He described the aid project which was intended to promote curriculum reform, how a new curriculum was actually developed, how he perceived teachers to be using text books and what official links there were between curriculum and assessment.

## **3.8 Observation**

The following observations were made:

- Classroom observations
- Observations of examination invigilation
- Observations of the marking of examinations.

### **3.8.1 Classroom observations**

My aim for classroom observation was to understand the practice of teaching in relation to assessment in the case study school: I aimed to gather examples of aspects of classroom teaching that I had learned about in interview; also to build up a picture of the lesson in its entirety. I observed all the mathematics classes of 45 minutes each, in

two grade 4 classes, two grade 3 classes and one grade 2 class, over a period of three weeks leading up to the mid year examinations of January 1997.

These observations were as follows. Fewer observations could be made in grade 3 than grade 4, since the pupils were sometimes practising for a performance when they would normally have a mathematics lesson.

**Table 3.1 Observations of classes**

<b>Grade</b>	<b>Number of observations</b>	<b>Initials of teachers</b>
Grade 2	18	SMY
Grade 3	28	NMH, HMA
Grade 4	36	HML, SMA

Participant observation was not an option in the classroom, in particular because the foreigner is so distinctive in Egypt. For this reason, non-participation and detached observation were used, as has been generally the norm in qualitative studies of education in non-western cultures (Hopkin, 1992, p.135). I was, however, participant to the limited degree that I did share some teaching of 'weak group' sessions with SMY. Although I did not record this experience in a formal way, I used it as contextual information for my non-participant, detached observations of teaching and its relevance to assessment.

During the classroom observation phase, an unstructured approach was deemed most appropriate for providing examples to back up interview data and for gaining a general picture of teaching as it related to assessment. The schedule is shown in Appendix A, which was used to prompt the observation and which provided headings under which to make notes in the class.

The schedule headings were derived from interview data as well as from the literature. The schedule was shared with teachers who were being observed so that they did not see me as an inspector of pupil performance during observation. But it was hard to persuade teachers they were not being inspected. The research diary of March 31st read:

It is hard being a non-participant observer: teachers are not objects. Both teachers I observed today of course wanted feedback and wanted to know what I had written. While mainly what I write is just descriptions, I must be making judgements too. I wish they didn't assume that us foreigners know better.

I usually sat at the front of the observed class, next to the teacher. I tried to look passively interested and not to interact with teacher or pupils, if possible, so as to gain as natural a picture as was feasible. There is no doubt, though, that pupils as well as teachers were intrigued, and sometimes excited or embarrassed, by my presence. One day I had a cold; as I took out a handkerchief to blow my nose, I realised that about 40 pairs of eyes were on me.

### **3.8.2 Observations of examination invigilation**

When I arrived at EGC, the headmistress said I could not be allowed into or near the examinations when they were taking place. At the time of the examinations, however, the attitude seemed less strict. On several occasions, I wandered around outside the spacious *legnas* (examination halls) as children worked; near the end I was allowed to take a photo in one of the *legnas*; and once I stood and chatted to one of the teachers as they invigilated. So I was able to make some informal observations of the system and these contributed to Chapter 5.

### **3.8.3 Observations of the marking of examinations**

One sense in which participant observation was possible was in the marking of examination papers. The First Teacher for mathematics asked me if I would mark the 600 grade 2 and 3 mathematics papers with all the grade 2 and grade 3 teachers, to make it quicker for them. This way I observed the process from inside and made notes later. Then the First Teacher for English asked me to help mark all the grade 2 and grade 3 English examinations. I also helped mark over 300 grade 4 English and French papers with the grade 4 and grade 5 teachers. These insights contributed to the content of Chapter 5 in particular.

## **3.9 Validation**

During the field research, the headmistress and two teachers read my account of the information the headmistresses had provided for Chapter 5 and commented on its accuracy. In June, 1998, SES, head of primary English, read and commented on Chapters 4, 6, 7 and 8. The headmistress received Chapter 7, which gave parents' views, to comment on if she wished although no comments were received.

For ethical reasons, teachers who knew the interviewees whose words were being reported, were not asked to comment on these accounts. The headmistress was not given pupil or teacher accounts, in case she could recognise individuals despite their disguised names. SES was in a good position to read all accounts since she had just left the staff of EGC by June 1998.

### **3.10 Issues concerning the research practice**

This chapter describes the methods used in the current research, including interview, observation, pupil writing and analysis of documentation.

The chapter shows that access was difficult, into an Egyptian state school and into the presence of Ministry officials. Access into the case study school was easy because of a personal connection with the school and the same was true for the pilot school. Access to pupils was not problematic, except that as the examinations grew nearer, those teachers who had not finished teaching the syllabus would not allow pupils to come out of their class for interview; these subjects tended to be English and mathematics.

Access to parents was only difficult if they did not wish to come into school because they felt embarrassed about their poor English, apart from in the case of one parent, Mira's mother, who refused to make any response when invited for interview. Access to teachers was superficially straightforward, but some teachers made themselves hard to find and others seemed afraid that the foreigner was actually a spy who would pass on everything they said to the headmistress. This fear reflected a subliminal hostility to foreigners generally; one reason for hostility was that foreigners were expected to be superior.

Some teachers resented being told rather than asked, on behalf of the headmistress, to be interviewed or observed, which perhaps made them more suspicious of the interviewer-observer. The teachers' reluctance raised the ethical issue of whether they should have had a real choice about taking part or not. Another ethical issue was whether research should be done which is dependent on people's goodwill, and for which they received no rewards or even benefits.

Other feelings connected to the research which the research diary highlighted, included the following:

- Frustration. This was because access was sometimes difficult and because tasks seemed to take a long time to get through. Because I depended on goodwill, it was impossible to force events to happen or complain when they did not.
- Tiredness. The research was particularly tiring because an extensive research programme had to be fitted into a short research period and was dependent on pupils, parents and teachers still coming into school before the summer vacation began. In addition, talking and understanding through Arabic as a foreign language was tiring. More than tiredness, the research diary described several bouts of illness, caught either from pupils or just because of the change of environment, and these illnesses did interfere with completing research to schedule.
- Anxiety about the computer. There were constant problems with computer technology because computers were still uncommon in Alexandria. This was why I had to reject the idea of using the NUD.IST software package for analysis. Printing was a particular problem and my options were an office in Cairo, which was four hours' journey away, paying a lot of money at the British Council in Alexandria or using a very old, poor quality printer which someone lent. In everyday Alexandrian life, the electricity was sometimes cut off suddenly so work on hard disc was easily lost; worse, the apartment where I stayed was vulnerable to intruders, to whom my laptop computer would have been very desirable. So I started carrying floppy discs around with me at all times. Because of risks like these, I had not been able to insure the laptop in England. Finally, the desert sand which forced its way into the apartment, and so the computer, was a constant worry.
- Enjoyment. The research diary contains numerous references to enjoyment, despite the more negative feelings described above. For example, on April 4th, 1996, the research diary read:

I am really enjoying myself immensely. It's strange, perhaps, that such mundane matters make me happy, but they do. I feel I have an idyllic lifestyle.

## **PART II. The research findings**

### **Chapter 4. The roles of assessment in Egypt from 1798 to 1996**

#### **4.1 Introduction**

This chapter analyses the aims of the government in Egypt and how the government used assessment to fulfil these. The chapter therefore looks more at the intended role for assessment and the system developed to implement this role, than at the effects of assessment in the lives of Egyptian individuals. While Chapter 5 looks at assessment within the context of the case study school as a whole, this chapter looks at assessment within the context of Egypt as a whole. Some effects and outcomes, in relation to individuals within the Egyptian case study school, are bound to enter the description in this chapter and in Chapter 5; but Chapters 6, 7 and 8 specifically analyse pupils', parents' and teachers' positive and negative views about the system as they experienced it at primary school level in 1996. Those chapters also describe national reforms to the system as they perceived them at primary school level.

This chapter traces shifts over time in government assessment aims and their implementation, drawing on historical data. It outlines initially policies of past Egyptian leaders which led the assessment system to take its recent form and shows how some political intentions for assessment have kept re-emerging over the past two centuries. The current aims of President Mubarak are then shown to focus on improvement of quantity and quality in education, especially primary education, and including the improvement of the individual's learning. The condition of the whole primary education system in 1996 is described, which puts in context the way the assessment system was operating. A picture of the shortcomings of both education generally and assessment specifically also helps explain President Mubarak's aims. His aims provide a rationale for recent reforms that sought to improve the learning of the individual as well as improve the system which assessed it, and improve the country which managed it.

## 4.2 The historical role of assessment in Egypt

### 4.2.1 The government's earliest aims for assessment

Napoleon invaded Egypt in 1798 and, although the French occupation only lasted three years, his influence on assessment policy was still discernible nearly two centuries later in 1996. To maintain power in Egypt, Napoleon needed a strong army and for this he needed educated men. The best candidates were selected and trained in a dedicated college for the army. When Napoleon left, Egypt's new leader, Muhamed Ali, adopted Napoleon's educational aims on behalf of the Ottoman Turks, who still had nominal dominion of Egypt. Muhamed Ali could perceive a relationship between provision of western-style (French) education and military, or political, supremacy. Muhamed Ali therefore set up secondary schools to feed into the military colleges and these elite schools provided food, clothes and a salary for their pupils, and brought teachers from Europe. A westernised education became a means to an official appointment to a public office with prospects of riches and status. This was strengthened after the introduction of an official qualifying certificate (Van Dam, 1983). This official qualifying certificate, based on Napoleon's new *Baccalaureate* examination of 1806, introduced the notion of central, state control of education (Broadfoot, 1996a). Most pupils who entered the elite schools and took this Egyptian qualifying examination were ruling-class Turks or other privileged foreigners. Egyptians continued to be educated in the religious Quttabs where children learned to recite the Qur'an by heart, sometimes for vocational purposes of becoming 'reciters' in rich houses. The government system was criticised in 1849, when Muhamed Ali died as being too closely, yet superficially, modelled on an inappropriate French system (Van Dam, 1983).

The earliest ancestor of the modern general secondary certificate (*thanaawiya aama*) was therefore the means by which only the most competent and also privileged in society were selected to defend and lead the Egyptians. The British superseded the Turks as occupiers of Egypt from 1882 to 1952 and their aim was broadly similar. Now, however, those who qualified entered the Civil Service rather than the military; and some Egyptians, albeit a westernised elite of them, could enter the colleges. The British assessed and selected those who would be of most political use to them as they ruled the state.

#### 4.2.2 The instrumental reasons for the first primary schools

*Primary* schools were established to prepare children to enter the secondary schools, but they charged fees so that only the better off could attend them (Cochran, 1986). In secondary school, numbers of pupils accepted after taking a competitive primary leaving certificate examination changed according to numbers of jobs that needed doing on graduation, but the choice of subjects taught did not adapt accordingly. The curriculum, being western and therefore secular, had no national models to follow with the result that the syllabus in both primary and secondary lacked planned objectives in terms of learning as well as vocation. So it became overloaded to include a wide but uncoordinated range of subjects. The aim of primary assessment under the British was, therefore, certification and selection of some more privileged Egyptians to enter secondary school where they could compete for qualification as employees of the British administration. Primary schools functioned as ‘... the only place one could go to get the necessary application forms for entry into the bridge-head zone, the only place one could go to learn how to fill them out’ (Dore, 1976, p.3).

The British, primarily Lord Cromer (Evelyn Baring), also encouraged a system of *elementary* schools whose planned objective was to train pupils for urban occupations useful to the British. However, access from the elementary school to the secondary school (and so college) was barred since a French and English examination had to be passed for selection to secondary, and these languages were not taught in the elementary schools.

Although in 1923, Taha Hussein, Minister of Education, made elementary schools free and compulsory, by 1930, only 18 per cent of the primary population were estimated to be enrolled in elementary or primary schools, although others attended private, often foreign, schools and, of course, the Islamic Quttab (Cochran, 1986). This was partly because rural families needed their offspring to work on the land, since there was no system of security for rural workers. Generally speaking, though, primary education was not considered important, except in light of future vocation. So when the Free Officers declared Egypt to be free of British oppression and of the Egyptian monarchy, in the revolution of 1952, national literacy was estimated at 30 per cent (Van Dam, 1983), the education system was fragmented and facilities were in poor repair. Only the



privileged could even enter the system by which prestigious government jobs could be reached through selective examinations; while some others worked for foreign or private trading companies; but most worked on the land.

#### 4.2.3 Revolutionary aims for assessment

The revolution of 1952 ushered in great social and economic change:

The aim of the revolution was to transform society. Education in the broadest sense was the fundamental means of achieving this... In the long term, the transformation would begin with the youngest child in the elementary school who would be educated in the principles of socialism, Arabism and national consciousness. The student would be taught (the word used was 'directed') to co-operate with the state in achieving its goals, it was hoped without question... It was an authoritarian approach, with the state the fountain of all goodness and wisdom (Hopwood, 1985, pp.136-7).

Egypt was unified by the revolution, under Gamal abd el Nasr who headed a central military government. The primary school system was unified in 1953 when elementary schools were abolished and primary schools made free and compulsory for every Egyptian child between six and 12 years of age. The curriculum was standardised and unified through the sole use of Arabic and through the national primary leaving certificate. Primary schools were further unified when foreign schools were abolished in 1956 (except Armenian and Vatican). The timing of examinations was a further unifying factor, with the whole country aware of national examination all taking place at the same time (Abdulla, 1985).

While using primary education as a means to disseminate the ideals of socialism, Arabism and nationalism, the new regime also made primary education accessible to all Egyptians. Unlike under the old regime, all Egyptians now stood the chance of selection to a higher stage of education by passing the primary leaving certificate or *ibtidaa'iya aama*, entering the new 'preparatory' stage or *adaadiya*, which could then lead them over the hurdle of the *thanaawiya aama*, or general secondary leaving certificate, and so into government jobs. These jobs had once been the reserve of a social elite consisting of foreigners, monarchy and landowners, all of whom had now been disempowered or expelled. In this sense, assessment was instrumental in bringing about social change and 'ordinary' Egyptians flocked into education in the hope of upward social mobility. For many students, '... education was their only passport to a

better future and they were deeply concerned with the likely effect of economic conditions upon their future prospects' (Abdulla, 1985, p.36.) In 1951, university enrolment was 35,000 but by 1961, it had reached 140,000 (Hopwood, 1985).

#### **4.2.4 Tensions between Nasr's political ideology and the economy**

As well as political supremacy, revolutionary ideals had included an unfamiliar focus on industrialisation, centralised economic planning, and rapid technological development (Faksh, 1976). In the National Charter of 1962, Nasr described science as the weapon with which revolutionary triumph would be achieved (Cochran, 1986). He therefore used education and its assessment as an instrument to improve the country's economy. But while all secondary school graduates were guaranteed government jobs, students who graduated in scientific subjects were more likely to become technocratic government advisers with more opportunities and higher pay than those who graduated from the arts who took whatever administrative job the government bureaucracy offered them.

For both sets of students, the aim of education was indirectly service to the state. The improvement the revolutionaries sought was national and political improvement. This meant that the state had control over employment and therefore social mobility was not guided entirely by choice. Immediate government employment following graduation, which soon meant graduation from university, began to be delayed because supply and demand had not been matched. Students studied at university the subject their *thanaawiya aama* grades qualified them to study. They accepted the fixed meagre salary in exchange for the security of a government job so that many graduates ended up in inappropriate jobs (Palmer *et al*, 1988). In these ways, the system seemed to favour government aims over those chosen by individuals.

It may seem surprising, then, that education continued to be highly valued by 'ordinary' Egyptians. This continued value is partly explained in that Nasr used an administrative or managerial approach to disguise some of the tensions between the ideologies of improved educational opportunities and national and political improvement. He used the administrative bureaucracy as the instrument of political change. In the same way, the objective nature of the *thanaawiya aama* (general secondary leaving certificate) and its well entrenched history as a legitimate certificatory and selective device helped

convince aspiring Egyptian people of the benefits and legitimacy of the system. The system did, however, come to be criticised. In 1968 and 1972 *redoublement*, following failure in end of year examinations, was introduced to grades 4 and 2 respectively to replace automatic promotion based on regular attendance. This was criticised as against the socialist ideal and, further, it ushered in the divisive practice of expensive private tutoring (Hanushek and Levy, 1994).

The *thanaawiya aama* was also criticised as being socially divisive:

The prevailing system of examinations ends in effect with a kind of class distinction since the decisive factor in passing such examinations is not real talent and aptitude but rather financial capacity and social environmental milieu (Abdulla, 1985, p.109).

#### **4.2.5 Sadat's shift in aims for assessment**

When Nasr was succeeded by Anwar Sadat in 1970, Sadat suggested

... to get rid of this overwhelming disease whereby many consider education as the instrument for acquiring a special social privilege, while the principal target for some educational people has become office-jobs irrespective of their value in the movement of society (Hyde, 1978, p.21).

In other words, Sadat claimed that the *egalitarian* ideal was socially divisive and it was also socially stagnating. Nasr's success in improving social opportunities through objective examinations was therefore only partial. Sadat now concentrated more openly on economics, recognising that both social and political aims depended on these. He sought the economic improvement of both the country as a whole and the education sector more specifically through the reintroduction of foreign investors. The economic devastation brought about by Nasr's political idealism had the effect of turning his nationalistic ideals on their heads.

The education system now came to reflect a mixed economy and divided culture once again. Private language schools were re-established where graduation was easier and links to lucrative private firms likely. Not only was private enterprise now encouraged but foreigners could run companies which offered salaries about twice as good as local private firms (Bartsch, 1995). Meanwhile, graduates from government schools took low paid jobs and went where they were sent. The assessment mechanisms put in place by Nasr to meet his aims of national social conformity were, therefore used instead, by Sadat, to meet his aims of national economic improvement. This was the situation when Hosni Mubarak became president in 1981.

#### 4.2.6 Mubarak's aims

In 1981, Mubarak faced a country in economic crisis, despite Sadat's attempts to improve the economy. Unemployment was increasing, and this was partly because the bureaucracy inspired by Nasr and maintained by Sadat was more than 12 times its original size in 1952 (Palmer *et al*, 1988). The bureaucracy was simultaneously draining the country of money and the population had reached 40 million in 1980, compared to 20 million in 1960. Mubarak combined social, economic and political aims. His speech of 1989 identified three national priorities: economic improvement, strengthening democracy and achieving educational reform. Education, especially primary education, was seen as a means both to achieving economic improvement and strengthening democracy.

In 1995, Mubarak's Minister of Education, Bahaa al Din, described educational reform as a matter of national economic survival (MoE, 1995, p.77). In a vein reminiscent of both Muhamed Ali and Nasr, he considered the wars of the future to be competitions in educational effectiveness:

War has become a war in mathematics, electronics and technology. This is a new concept of war to people. Consequently, military supremacy is the outcome of education (MoE, 1995, p.28).

Unlike Muhamed Ali, however, and more in line with Nasr, this aim embraced a dedication to universalising access to education at the Basic Education level, by the year 2000, and the eradication of illiteracy by 2002. These goals had been inspired by the Jomtien 'Education for All' conference of 1990 (Zaalouk, 1995). Unlike the 1952 revolutionaries, though, Mubarak's government emphasised that universal quality of educational provision was as important as comprehensive provision.

Like Nasr, Mubarak believed that economic soundness could be reached through education, by concentrating education in the fields of applied mathematics, science and technology. Like Sadat, he saw economic improvement through increasing dialogue with western nations, which itself necessitated emphasis on learning to communicate in foreign languages. These combined goals were again reminiscent of the earliest French initiatives at the start of the 19th century. In Mubarak's words:

Egypt is now witnessing a comprehensive renaissance in all aspects of life which aims at accelerating its potential to cope with advanced countries. This cannot be achieved without the assimilation of the orientations of the modern world and its

new patterns... the countries which possess knowledge and technology will have the right to survive, and those who don't will lose their way and suffer from ignorance and poverty... (MoE, 1995, p.78).

Since education was a matter of national survival, its responsibility no longer fell only to the experts but was to be democratically distributed. Mubarak described education as a 'national mission in which all authorities, legitimate channels and individuals' had to take part. Since it was to reflect public opinion and aspiration, its basis was now (liberal) democracy (MoE, 1995, p.34). This emphasis on democratisation resulted from a desperate need for the whole nation to unite to improve education. The national unity indicated here differed from the national unity stressed by 1952 revolutionaries, in that revolutionary policy emphasised central prescription as the agent of unity while Mubarak implied democratic participation. Nasr stressed the equality of Egyptian citizens while Mubarak suggested their individual potential.

When Mubarak came to office, the potential of assessment for educational improvement was specifically recognised for the first time. In 1990, the first National Centre for Examinations and Educational Evaluation (NCEEE) was established under the then Minister of Education, Fathi Sorour. The aims of this independent body were that through assessment, the system of education could be improved and so could the quality of individual learning, through the training of teachers, inspectors and administrators in relevant aspects of assessment; through the improved efficiency of test development; through research related to examination reform. The aim was to improve assessment, which in turn would improve the system and improve individual learning, which in turn would improve economic development as well as contribute to democratisation.

## 4.3 The current system in Egypt

### 4.3.1 Primary education provision

The primary education system that President Mubarak aimed to improve is outlined in the following paragraphs. Table 4.1 shows how, at the time of this research, pre university education in Egypt was divided into three phases, primary, preparatory and secondary. Primary and preparatory school was free and compulsory for all children and constituted Basic Education. Enrolment in Basic Education was officially over 90 per cent for boys and girls, although in reality it was probably between 70 and 80 per cent (Carroll, 1995). Grade repetition increased from 9 per cent in grade 2 to 22 per cent in grade 5 (MoE, 1993b, p.13). Pupils dropped off in transition to the preparatory stage and repetition of the first year of preparatory school was about 12 per cent for girls and 14 per cent for boys (World Bank, 1996). Kindergarten provision was increasing with nearly 300,000 children attending, but as yet it remained almost exclusively in the private sector and does not therefore appear on Table 4.1. As shown on the table, a small minority of all schools were private.

**Table 4.1 Numbers of schools and pupils at different school phases in Egypt, 1995-6 (MoE, 1996)**

School phase	Schools	Pupils
PRIMARY grades 1-5 <i>ibtidaa'iya</i>	16,188	7,470,437
Government	15,058	6,937,034
Private	1,130	533,403
PREPARATORY grades 6-8 <i>adaadiya</i>	6,496	3,409,127
Government	5,910	3,292,006
Private	586	117,121
SECONDARY grades 9-11 <i>thanaawiya</i>	2,995	2,602,811
Government	2,435	2,402,281
Private	560	200,530

The education system was divided into 26 administrative governorates, the largest being Cairo with over 800,000 primary school pupils. Although the number of pupils seems huge, the number of schools was proportionately very small, reflecting the use of double or treble shifts, whereby one building effectively housed three schools with three

administrative systems, each day. There were fewer than 5,000 primary schools offering a full day compared to more than 11,000 who offered shifts (MoE, 1996). Children in multi-shift schools were calculated to lose five hours per week of instructional time and nearly all extra curricula activity. The buildings in which primary schools were housed were also often in bad repair and in addition, shortages of space led to large class sizes, frequently reaching 60 or 70. Access to schools was difficult in rural areas where communications were poor, a situation that inspired UNICEF's thriving community schools programme for villages (Zaalouk, 1995).

#### **4.3.2 Teaching at primary level**

Teacher quality and commitment were considered generally low because most teachers were inadequately trained at the pre service stage, and received little in service training. Teachers were less than adequately paid because their salaries were in line with the whole of Egypt's enormous government bureaucracy (Cummings, 1986); a teacher's starting salary was around LE100 (£20) per month. Poor salaries encouraged private tuition which detracted energy away from the teacher in class, they also raised the cost of 'free' education by an estimated LE7 billion per year across pre university education which was about the same as the Ministry's annual education budget (World Bank, 1996). In addition, all dimensions of teacher compensation appeared unrelated to performance.

All teachers in even the most independent and wealthiest schools were accountable to government regulations for curriculum coverage and assessment. The national curriculum took the form of national textbooks provided to every child at the beginning of each school year. Material in these textbooks was generally considered to be pitched too high conceptually for each age group who received them [USAid curriculum trainer, interview notes]. These issues are dealt with more fully in Chapters 5 to 8.

#### **4.3.3 Selective examinations**

While the following section chiefly concerns the primary phase, an account of important assessment points at other phases is given, both to provide context to the primary level and because the attitude to assessment at secondary level influenced assessment at primary level. Examinations across all three phases tested pupils' knowledge of material in government textbooks which were compulsory in all classes in Egypt.

In 1996, the only assessments which were truly selective for a subsequent phase of education consisted of the following; each of them also served a certification role, but not in isolation from their selective role. At the end of the five years of primary education, in grade 5, children sat the *qabuul* examination, which was developed by the local governorate. This limited the choice of children in government schools as to which preparatory school they could attend for the final three years of Basic Education. This could be Commercial school, for the lowest achievers; Nursing school; Technical school; or General preparatory school, for the highest achievers.

In grade 8, at the end of the preparatory phase, pupils took the *adaadiya* examination developed by the local governorate, on the basis of which they received the *adaadiya aama* certificate; or the Commercial, Nursing or Technical Basic Education Certificate; or, if they failed altogether, a Certificate of Completion. In 1994, about 50 per cent of preparatory schools pupils entered employment directly (De Luca, 1994). On the basis of which certificate each pupil gained, they were selected either for academic general secondary school or for technical secondary school, while graduates of commercial or nursing schools went directly into employment. Students in technical secondary schools stood little chance of entering university and were expected to enter the employment market immediately on leaving secondary school. A good grade in the *adaadiya* examination in grade 8 was therefore indirectly necessary for progression to university. This was less the case in private schools where promotion from grade 8 to grade 9 was more or less automatic.

Those who proceeded to technical secondary school, mostly gained the Technical School Diploma and then entered the job market, or sat entrance tests for a Higher Institute of Technology. The *thanaawiya aama* was the critical selective tool for pupils leaving general secondary schools who were hoping to proceed to university. The *thanaawiya aama* was set and marked at national level. The current and historical importance of the *thanaawiya aama* examination has already been suggested. In 1996, it still provided access to a university place and, according to each pupil's final grade, determined which faculty they could enter within the university. As competition for places at university increased in line with population growth, grades for priority subjects, which continued to be medicine and engineering, constantly increased to as high as 92 per cent in 1996.



A radical scheme to separate university entrance qualifications from school leaving qualifications had failed to come to fruition (Sorour, 1990). Minister of Education Fathi Sorour had fought to increase the certification function of the *thanaawiya aama* and make the selection function more effective for prediction. He proposed that the certificate awarded for the leaving examination provide employers with useful descriptive information rather than just a grade; and also that, to enter university, dedicated entrance examinations should be taken. In addition, he proposed that items within both sets of examinations be improved to include assessment of critical thinking, problem solving and creativity.

Instead, a revised version of the old system of the leaving certificate was introduced when Sorour was succeeded by Bahaa al Din. He did not increase the certification function of the *thanaawiya aama* and declined to change the quality of item (Carroll, 1995). What could have been assessment for certification and improvement of the learning of the individual therefore remained a selective examination. He did, however, give pupils more choice in the subjects they studied and initially allowed them more chances to sit the examination, although this was later reversed again.

These reforms to the *thanaawiya aama* were based on the International General Certificate of Secondary Education (IGCSE) system which was reallowed in Egypt in 1990. Five per cent of all applicants to the university could gain entry via the IGCSE examination. In 1996, this meant that those who could afford the high fees to pay for an IGCSE course stood a better chance of entry into university. The same was true for the International Baccalaureate. These western courses were also considered less stressful than the *thanaawiya aama* [SES, interview notes].

A shift in emphasis, which accompanied both ministers' revisions, was on the division between qualification and work. Nasr's promise of work for all graduates was no longer a guarantee under Mubarak: he stressed the need for graduates to take their own initiatives to create work (MoE, 1992). The high status of being a graduate, however, did not seem to depend on this guarantee, and children still saw selection in the grade 5 *qabuul* examination as the first step towards success in the ultimate *thanaawiya aama*. Like Napoleon's earliest version of this examination, it allowed them into one kind of elite that non-graduates could not enter.

#### **4.3.4 Non selective examinations**

Examinations written by the individual school were taken compulsorily, twice a year by all children in non certificate grades except grade 1. During grade 3, children sat a new external 'certificate' examination written by the local zone, which they had to pass to be promoted to grade 4 with their peers. Neither this certificate, nor the school written examinations were actually selective; but neither did they play a pure certification role. Even failure in the primary school certificate, the *qabuul*, only meant repeating grade 5 once, since education had been free and compulsory up to the end of grade 8 since 1981. These examinations simply provided stumbling blocks to those who could not keep up at the speed demanded by the curriculum.

### **4.4 Improvement of the individual's learning**

#### **4.4.1 Curriculum reform to improve the learning of the individual**

For educational reform to be effective in contributing to the country's economic and political development, it was clear that a recently increased enrolment in Egypt had to be matched by improved quality in the processes of the learning of individuals, not only in the results of their examinations. Some of Egypt's economic and political problems were attributed to children's inability to take an active individual role in their own learning, to apply knowledge, to solve new problems and to think creatively as individuals (Sorour, 1990). This encouraged a new emphasis on processes of learning as well as outcomes.

In 1993, a Conference on the Development of Primary Education Curricula (MoE, 1993a), chaired by President Mubarak's wife Suzanne, discussed reforms which would move children away from rote memorisation and passive learning through teacher transmission, towards this model of active individual learning (MoE, 1995). As was traditional in Egypt, reform of learning was implemented primarily through curriculum improvement which meant through the revision of textbooks and then the assessment of textbook material. The sixth grade of primary schools had just been abolished to release funds for improving the quality of learning in the other five primary grades. Holding a perspective which assumed certain *basics* to be essential, the conference

therefore recommended that grades 1 to 5 be divided into two phases. In grades 1 to 3, the curriculum would steer children towards acquisition of basic skills in Arabic, mathematics and religion and so ensure a sort of 'minimum competency' (Bracey, 1983). In grades 4 and 5, the curriculum would encourage children to apply these basic skills and substantiate their knowledge through further subjects, including a foreign language and science. Especially in the lower grades, the curriculum would be greatly reduced to allow depth of learning as well as breadth. These changes indicated a challenge to the traditional *transmission* model of learning; and with this challenge came an increased emphasis on the processes of learning.

#### **4.4.2 Assessment to improve the learning of the individual**

Curriculum, textbook and assessment were intimately linked in that pupil assessment traditionally consisted primarily of written tests of material rote learnt from the government textbook. To implement curriculum and textbook reform in order to improve individual learning, assessment therefore had to be adapted accordingly. The Conference decided that an externally developed *certificate* examination should be administered during grade 3 and grade 5. Since the goals of the new curriculum were innovative in their movement away from rote memorisation and passive learning, these assessments were seen as a powerful means of causing the new emphases to be observed in practice. The NCEEE, who became responsible for developing the grade 3 and 5 examinations aimed to develop these new examinations in discussion with representatives from the 26 governorates and then to implement them nationally so that basic minimum standards could be monitored nationally and individual learning could be emphasised, including the learning of how to apply skills and knowledge and how to use initiative and problem solving skills.

The MoE was afraid that the new examinations would increase the number of children who had to repeat the grade and so worsen the economics of education. If the examinations included innovative elements which could not so easily be crammed for as the old rote memory tests, the Ministry feared that the quality of teaching would need attention, first. The NCEEE therefore developed a suite of diagnostic tests in Arabic and mathematics for both grade 3 and grade 5, so that remedial teaching could take place before pupils attempted the external examinations, or at least before the repeat

examination at the start of the next grade. A diagnostic approach in assessment was a further innovation.

These reforms were not to be. The 26 governorates refused to co-operate with the NCEEE over development of the grade 3 and 5 examinations so that no national system could be implemented and each governorate continued to develop their own, idiosyncratic assessments which would not necessarily reflect adequately the new emphases on learning processes which the new curriculum embraced (MoE, 1993b; Carroll, 1995).

The Ministry, in addition, took no action to promote the diagnostic tests, considering their approach too alien to the marks-oriented culture of Egyptian testing [First Undersecretary for General Education, interview notes]. Instead, the Ministry lowered the pass mark for the examinations, which meant that children could pass even without answering correctly any items that assessed higher order skills; and the move away from results and towards processes, was not supported.

The new examinations were developed individually by each governorate and were first administered in 1994. The NCEEE's task now was to develop means of making the examinations as standardised as possible across the country and as innovative as possible, in terms of including items which assessed the higher order skills, despite the resistance they had met. In 1995, therefore, centrally prepared examination specifications were made compulsory for every primary grade and for every governorate. These specified the format of each examination, including how many items should be included to assess more than rote memorisation skills. After each session of examinations, the governorate was to collect a sample of every examination administered under its auspices and send it the NCEEE for their monitoring. Although the immediate effectiveness of these measures has been questioned (De Luca, 1994; MoE, 1993b), their aim to improve quality in individual learning must be recognised as a long term goal, which would be reached gradually.

The NCEEE, further, arranged the training of 12,000 grade 3 inspectors between 1993 and 1996 in the writing of valid assessments. These inspectors then took responsibility for training the First Teachers whom they inspected in schools and who wrote in-school

examinations for grades 2 and 4. Training of grade 5 inspectors was to follow [Head of training and dissemination at NCEEE, interview notes].

The NCEEE also developed and disseminated guides or *taqwim* (an Arabic word meaning 'strengthening'.) The *taqwim* guides for 1996 to 1997 were more dynamic than each preceding version, in terms of focussing on the processes of learning. These latest editions at last contained a definition of the objective of each unit within the textbook so that individuals could identify learning intentions within the curriculum. They also outlined the skills children needed to bring with them from previous experience. They then provided example-items to assess each topic within the text book, giving a marking key with various possible answers rather than the traditional 'one model answer' list and its implication that there was one 'right' answer. Finally, they offered example diagnostic tests relating to three units of each text book at a time. Through the *taqwim* guides, the NCEEE aimed to exemplify practices they had failed to implement compulsorily, but the success of this approach depended on teachers and pupils receiving and making use of the guides for themselves.

#### **4.5 Assessment for accountability and system improvement**

All schools followed the same syllabus at the same time and were accountable to government for this. All schools in Egypt were accountable to the government through the submission of bi-annual examination grades and through frequent government inspections, on the basis of this syllabus. As described above, accountability to the state was a keystone of the 1952 revolution, and education had been a primary political tool since before then. While examinations served an important certificatory and selective role for pupils, parents and teachers, submission of their results fulfilled an accountability purpose for government along with reports from inspections, which put government into a powerful position over individuals in schools. Until secondary leaving certificate level, however, schools were accountable only at local, governorate level; governorate representatives were then accountable to central government.

Monitoring of national standards at primary level has not been customary in Egyptian education, despite the competitive atmosphere attached to examination results and their

use for accountability of individual schools and individual teachers. De Luca wrote about central administration in Egypt, that ‘... overlap in roles and responsibilities ... means accountability is somewhat diffuse, and information flow, while improving, is still limited’ (1994, p.41). National monitoring had been expected to ensue from the centralised grade 3 and 5 examinations; but the governorates boycotted centralised administration to avoid a comparison of standards. A survey was, however, administered in 1993 to measure equivalence across governorates. The report of the survey recognised that

... despite the official examination construction specifications that are carefully prepared at the national level, different educational idaras [local areas] still manage to construct examinations that are not equivalent... (MoE, 1993b).

Although the NCEEE provided recent examination development training for a sample of grade 3 inspectors, the quality of item-writing continued to be variable across local areas, too, and there continued to be ‘...considerable variation in the content covered in different forms of the same examination...’ (De Luca, 1994, p.39). The NCEEE therefore suggested to the government a programme of national sampling whereby two or three governorates each year requested that the NCEEE write, administer and monitor their grade 3 and 5 examinations on a voluntary basis. Over a decade, substantial national monitoring data could be built up. By 1996, two governorates had agreed to allow NCEEE to develop and monitor their external examinations [Head of NCEEE examinations department, personal interview].

The MoE also supported the NCEEE’s development of a national inspectorate on the lines of the English Office for Standards in Education. Their role was to look at practical as well as academic standards on an *ad hoc* inspection basis, making their judgements more objective than was currently the case. By the summer of 1996 stalemate had occurred in this scheme since the Ministry could not supply the necessary funds for an effective national inspectorate.

#### **4.6 Overview: the roles of assessment in Egypt from 1798 to 1996**

Egypt's leaders since Napoleon in 1798 had political, economic and social aims for assessment. They aimed to strengthen Egypt politically by improving its military power and so its economic power. This necessitated social shifts too. The political, economic and social nature of assessment was epitomised by Nasr's introduction during the revolution of one selective, written examination called the *thanaawiya aama*, by which anyone could reach socially and politically desirable positions, especially if they achieved good results in science and technology rather than in the arts.

These political, economic and social aims initially had the effect of unifying the country, which allowed for strong central control and therefore allowed the government to disseminate government ideology effectively. Although social shifts soon turned Egypt into a divided but democratic state rather than an egalitarian socialist state, the strong central control remained in most fields, including education. Strong central control also led to the need for the central government to legitimise its monopoly, and the 'objective' selection examinations for school leavers at each school phase provided this legitimisation. In Broadfoot's words, though, '... an apparently objective, scientific assessment is perhaps the most effective form of educational (and thus social?) control yet developed... an appearance of conformity with rational norms' (1996a, p.144). Having legitimised the means of assessment, the central government could demand that those involved with assessment show accountability, both through examination marks and using government inspectors.

The effective role of such a system of selective examinations and accountability to a central control was to inhibit the government's ultimate aims of political, economic and social improvement. Economic aims were thwarted because central, bureaucratic administration was expensive and unwieldy. Social change was inhibited for all but one elite section of society because accountability and the unchallengeable rationality of the examination system left most people unable to act freely: they could choose to strive towards a government position but they could not choose how to strive for it or which position to strive for. If Mubarak's recent emphasis on democratisation had made more impact, the 26 governorates may have agreed to co-operate in the democratic drive to

improve the quality of individual learning through national assessment at grades 3 and 5, but instead the governorates resisted further accountability to central control.

The assessment goals of selection and accountability also made formal, written examinations, based on rote memorisation, the easiest instrument of assessment, especially since pupil outcomes from examinations were numerical and so were straightforward to use for selective and accountability purposes. Examinations traditionally motivated children to learn by heart in order to achieve high marks in written examinations. The aim of these examinations had never been for individual learning in any other sense. Their aim had been the instrumental aim of the government to select a useful elite; and the aim of the people was to be selected into that ruling elite. Whatever processes had to be endured by individuals on the way towards this objective were subsidiary to the objective itself.

Recently, individual learning that does not rest only on memorisation has been seen to be instrumental in achieving the government's ultimate aims of political and economic improvement. Assessment of the processes of individuals' learning, now in the sense of their deep-level knowledge and understanding (Crooks, 1988) which children can apply and extend, has therefore recently received central government attention.

The NCEEE has for this reason specifically introduced improvement to primary (and preparatory) examinations themselves, to include assessment of these higher order skills. The aim of these improvements to assessment was improved individual learning. The MoE and its representatives in the governorates, though, were not always in agreement or even communication, with NCEEE. Their drive for improvement did not, therefore, represent the common '... shift in the role and orientation of assessment from selection to enhancement of learning, and from a preoccupation with reliability to an emphasis on validity...' (Linn, 1997, p.316). Rather, it was through improved individual learning that children could now achieve the highest marks in improved examinations and so pass through the *thanaawiya aama* and so, indirectly, improve the political and economic prospects of Egypt as well as their own social prospects.



## **Chapter 5. The roles of assessment in the case study school**

### **5.1 Introduction**

This chapter analyses the roles of assessment in El Nasr Girls College Primary School, the case study primary school. The broadest definition of assessment is used in this analysis, as given in Chapter 1. In Chapter 1, assessment is defined as the gathering of information; measurement of students or institutions; evaluation or judgement about achievement; and as the means or instrument of testing. This chapter attends more to the purpose and implementation of policy than to its effects; more to practical assessment arrangements than to people's views about them (see Chapter 3). It addresses, in particular, the question, what were assessment's official roles in the case study school? Chapters 6 to 8, in contrast, draw on data from interview with pupils, parents and teachers respectively and analyse these parties' views as to the effects of assessment on individuals.

In Chapter 2, three interconnected official purposes were described for educational assessment: certification and selection; accountability and system improvement; and the improvement of individual learning. In this chapter which focuses on the official roles of assessment, the means and purpose of assessment in the case study school, El Nasr Girls College (Junior), are analysed in terms of the three categories. To put discussion about assessment into context, a description is given first of the nature, history and structure of El Nasr Girls College (EGC).

### **5.2 The case study school**

#### **5.2.1 The case study school in the context of national school types**

El Nasr Girls College was set up by the English in 1935 as the English Girls' College during the period of British protectorate in Egypt. It provided an English style education using the English language for girls of elite families in Alexandria and from further afield in the Middle East. Staff were sent from England and, from 1945, lived on the school campus. In the 1952 revolution, the Free Officers of the Egyptian military overthrew the Egyptian monarchy and disentangled Egypt from English

domination (see Chapter 4). Nationalisation was an essential policy in the government of Gamal Abd al Nasr which followed. In 1956, therefore, all foreigners, unless legally espoused to Egyptians, were evacuated and property held by foreigners confiscated. EGC was taken into direct government control before becoming one of a group of National Institutes. These held special status as foreign language schools alongside government founded Arabic schools; they were owned by a board of governors who had close government links. There were nine National Institutes in Alexandria, and El Nasr Girls College (EGC) was one of these.

The popularity of these *language* schools and the insufficient number of Arabic schools led during the 1980s and 1990s to an expansion of language schools. The newest establishments, however, were completely independent, owned by a single entrepreneur and only supervised by the MoE. The existence of private schools was initially considered antithetical to Nasr's philosophy, which was why the oldest private schools were nationalised; but as need dominated over ideals, new private schools were tolerated and a new era of private enterprise was initiated (see Chapter 4).

The case study school, EGC, could therefore be considered as bridging the gap between government Arabic schools and private language schools in that, while it was originally founded for independent reasons, the MoE developed a penetrating control in its organisation after 1956. As an English medium school, EGC taught mathematics and science in English; but the syllabus for these was translated from the government Arabic text and the content for all subjects except English itself was the same for all schools. *High level* English, however, was taught in EGC and in private schools from kindergarten level rather than starting *low level* English at grade 4 as did government Arabic schools.

Fees at EGC were higher than those in government Arabic schools of which most were practically free of charge; and yet substantially lower than in the new private language schools. Other than the fees for the British originated International General Certificate of Secondary Education, the highest year's fees at EGC were for senior 1 where they amounted to nearly LE1,100; this was equivalent to about £220 in April 1996. At lower primary level fees came to less than LE900 (£180). In one of the private language schools in Alexandria where the pilot study was carried out, by contrast

infants' fees came to about LE2,000 (£400), which was paid annually after a down payment of about LE5,000 (£1,000) on acceptance of a school place. In another private primary school, there was no down payment, but LE 1,500 (£300) was paid each half year.

The case study school's position between the two extremes of government and private education was reflected in many ways in its assessment procedures. The control exerted by the MoE despite EGC's semi private status was an especially prominent feature in these procedures. Despite the expansion of private language schools, they remained a minority of seven per cent at primary level nationally, with the great majority of schools in Egypt and in Alexandria being government Arabic schools. The numbers and proportions of both types of school in Egypt for the 1995 to 1996 scholastic year are shown in Table 6 in Appendix B.

### **5.2.2 Structure of the case study school**

EGC, at the time of research in the spring of 1996, was a school of nearly 5,000 pupils from five to seventeen years. A few boys, 34 in total, whose parents were staff at the school, remained in the lower school, but this system was being phased out to leave only girls. The school was divided into five departments. The five departments included the kindergarten, *hadaana*; the junior school, *ibtidaa'iya* the middle school, *adaadiya*; the senior school, *thanaawiya*; and the International General Certificate of Secondary Education department (IGCSE).

The kindergarten consisted of two year groups totalling just over 400 pupils aged between five and seven years who were housed in their own building. These school years were not compulsory, although most schools, and all private schools, provided for them and it was a Ministry aim to include them within the period of compulsory, free education in the near future (MoE, 1995). Currently, the compulsory stage of education, known as Basic Education, consisted of the first eight years of schooling, beginning with the first grade of primary.

The primary school, called the *junior* school at EGC, had five grades, reduced by the Ministry from six in 1988 and likely to be expanded to the original six years again in the near future (MoE, 1995). In every grade there were seven or eight parallel classes of at least 45 pupils each, making a total of about 300 children in each year group. Grade 1,

known in EGC as 'Junior 1', was in a separate building and was the only junior grade where a system of class teachers, rather than subject teachers, functioned. Grades 2 and 3 worked together in one building close to the grade 4 and 5 classes which used a separate one, and there was some spillover between the two.

The preparatory or junior secondary school, known at EGC as the *middle* school, functioned in another block, with pupils aged between eleven and fourteen years; three year groups were divided into about ten classes each. A similar arrangement existed for the upper secondary, or *senior* school. The senior department was, however, divided between a minority of pupils who had opted to study for the International General Certificate of Education as their final qualification, which was just over 100 pupils, and the majority who had opted to finalise with the Egyptian national equivalent, the *thanaawiya aama*, which consisted of over 1,000 girls. EGC was one of few schools who were then offering the option of an IGCSE qualification.

The numbers of pupils in each grade is shown in Table 7 of Appendix B. It is striking how many pupils were at each primary stage. The large numbers within each grade and then within each class had implications for assessment, increasing competitiveness and necessitating careful accountability monitoring as well as limiting options in assessment for individual learning. Although EGC tried to cling to its original 'family' atmosphere, the large numbers tended to preclude attention to individual needs.

### **5.3 The role of assessment for certification and selection in the case study school**

#### **5.3.1 Certification as curriculum coverage**

Certification was referred to in Chapter 2 as denoting a declaration or formal statement which detailed specific educational performance. Assessment for certification, it was observed, often shared a role with assessment for selection whereby children's suitability for a future course of study or work was predicted. In the case study school the two roles were so closely connected that the certification role was not apparent except as part of the selection process. Despite the new, external grade 3 and grade 5 examinations being deferentially named *shahaada* or certificate examinations, the certification they provided fed mainly into the selection process. No children needed certificates to enter employment at the end of the primary stage since all children were

continuing to middle school. Children were therefore certified chiefly as having covered earlier stages of the curriculum, since this allowed them to continue on to later stages.

The lack of distinction between certification and selection was paralleled by the close connection between national curriculum and assessment. The curriculum was designed with a view to skills which were relevant for primary aged children (personal interview with USAid curriculum trainer), but their relevance was overwhelmingly in the context of a hierarchical curriculum in which every stage had to be mastered, rather than of how children could apply skills and knowledge in their immediate lives. Concentration on mastering a stage, for example, learning how to 'perform' long multiplication in order to pass examinations for a given level, often meant that simple but easily applicable skills such as using number bonds to ten or times tables were weak because they were not specifically assessed at that stage. Certification was of covering specific units in the book.

Since curriculum coverage was what determined good test and examination marks which in turn indicated mastery of a given stage, there was little incentive to teach or certify skills or understanding which did not feature in the government curriculum. Such a limited view was encouraged also by government prescriptions which specifically prohibited deviation from the curriculum and by large numbers in class and limited time.

### **5.3.2 The means of assessment for certification and selection**

Assessment for selection and to some extent, certification, was mainly by written tests and examinations. The structure of their administration is depicted in Table 5.1 The table shows how the school year from September to May was dominated by written assessments at every stage through which children were expected to prove their ability. Examinations written by the school were taken twice a year, in January and May, by children in all non *shahaada*, or certificate, grades except grade 1. During grades 3 and 5, children sat two semester examinations of an external certificate which they had to pass to be promoted to the next grade with their peers. Those who failed any year's examination had another opportunity to be examined on the whole year's work in the *mulhaq* repeat examination which took place in September. By law, if children failed

the grade 3 or 5 examination at the repeat examination, they had to repeat the grade, but only once. By law, failing the school-written grade 2 or 4 repeat examination did not prevent a child from automatic promotion, although for motivation reasons the school sometimes encouraged the rumour that promotion was dependent on examination results in grades 2 and 4 as well as in grades 3 and 5.

Every month in which no examination was held, children in all grades from 2 to 5 sat a 'monthly test' in the non practical subjects. The monthly test assessed one unit of the prescribed government text book. The Ministry therefore decided how the year's syllabus was broken down into teaching and assessment chunks. The First Teacher, who had most experience with a particular age group and a particular subject, wrote the monthly test for all children in that subject and grade. All children of the grade took the test in the same week, but they sat in their usual classrooms with their usual teachers, who then marked the tests themselves; this was in contrast to the high security which regulated the bi-annual examinations and particularly the external examinations of grades 3 and 5.

The role of the monthly test had officially been shifted away from selection towards certification and learning. Previously, the marks from monthly tests were combined and contributed 20 per cent to each child's end of year results. The Ministry had, since 1992, abandoned this scheme since it was claimed that teachers could use these marks to encourage pupils to take private lessons with them; or to reward those who already did [First Undersecretary for General Education in Egypt, interview notes]. In effect, however, pupils saw monthly tests as a channel by which to prove themselves, but one which was less important than the examination channel. Despite their new role, teaching and learning were still oriented towards these written assessments *as if* they had an effect on final marks.

**Table 5.1 Structure of the administration of written assessments**

<b>Grade</b>	<b>Written assessment</b>	<b>Months</b>	<b>Subjects</b>
1	Monthly tests	Oct-Apr	Arabic, mathematics, religion, English
2	School-written examinations	Jan and May	All subjects
	Monthly tests	Oct, Nov, Feb, Mar	All subjects
3	External certificate examinations	Jan and May	Arabic, mathematics, handwriting, English
	School-written examinations	Jan and May	All other subjects
	Monthly tests	Oct, Nov, Feb, Mar	All subjects
4	School-written examinations	Jan and May	All subjects
	Monthly tests	Oct, Nov, Feb, Mar	All subjects
5	External certificate examinations	Jan and May	Arabic, mathematics, handwriting, religion, English, social studies, science
	School-written examinations	Jan and May	All other subjects
	Monthly tests	Oct, Nov, Feb, Mar	All subjects

Nearly every month, children received reports in which test or examination grades for all tested subjects were given and short comments made. These were written by the form teacher, who was typically an inexperienced, junior teacher; she took test results from the class mark sheets. Comments had prescribed meanings, for example, those with the highest marks took 'excellent', the next highest took 'very good' and so on. Results and reports on monthly tests therefore provided a continuous assessment of the child's progress in terms of syllabus coverage, since each monthly test assessed material from one unit of the prescribed text book. An analysis of exact skills or knowledge mastered or not mastered was not, however, made, except that the content of the syllabus was commonly known; most attention was certainly paid to marks rather than comments.

Assessment for achieving marks was by written tests and examinations in all but a few cases. Oral assessment had limited formal manifestation. In grade 4, the Ministry had just introduced English as a second language into all schools. It was felt that by mid year, children would not have learnt enough English to sit a written mid year examination, and therefore, an individual oral assessment was made. Children in grade 4 were also assessed orally in basic studies which was primarily a practical subject relating to craft, design and technology. This had also been newly introduced by the Ministry.

### **5.3.3 The purpose of assessment for certification and selection**

#### ***The selective function of assessment***

Promotion was not actually barred to those who failed the examinations. The norm referenced system in EGC, however, functioned *as if* the truly selective mechanisms of the government system were in force, which currently allocated places according to examination results at grades 6 and 9. It also functioned *as if* the now defunct tradition whereby pupils could be selected out of education at the end of primary education still existed, even though in 1981 education had become free and compulsory up to the end of grade 8.

Finally, it functioned *as if* the truly selective mechanism for entrance to university, the *thanaawiya aama* examination, was also the system at primary level. Belief was frequently expressed by parents and teachers that children's performance in the earliest grades would actually determine how they did when selection really did exclude those who performed badly (Hargreaves, 1997a). One teacher explained:

If children do badly, parents come to the teacher and cry. Teachers tell parents that children will develop as they get older, but parents believe that how they are in Junior 2 is how they will go on. So they push, because they want their child to become a doctor or an engineer, and for that they need high marks [SMY, interview notes].

In exceptional cases a child performing badly at primary level in EGC could be referred to an Arabic school, but in contrast to children in government schools, most children continued through to meet their first real selective barrier on entry to university (see Chapter 4).



Selection in the case study school was therefore chiefly vicarious. But it was also genuinely selective in that children were 'chosen in preference to others' for being of 'particular quality' (McLeod, 1987). Assessment selected children within the class against each other and against an authoritative standard of full marks.

### *Competition to achieve high marks*

Competition for high marks in class mirrored the now non-existent race for places in school. The norm against which all children were assessed was full marks in all the tests and examinations *as if* they were truly selective; or against the highest mark achieved in the class. Rejection could be sensed by children who were not among the top scorers in the tests and examinations *as if* they had not been 'selected'. They felt the need to prove themselves through their marks.

This selection-like competition was observed in most lessons. Sometimes, direct comparisons were made between weaker and stronger children. In front of one child, the teacher explained aloud, 'This girl is very weak' [SEY, observation]. The children themselves exposed weaker neighbours; for example, comments like 'Miss, Yasmina doesn't understand' were frequent [HMA, observation]. Children entered EGC at different ages so that in any one class an age range of three years was common. Yet the youngest and oldest competed for top marks in the same bi-annual examinations and monthly tests. In class, all children were therefore expected to work at the same, challenging speed. One geometry teacher [SMA, observation] was heard to say, 'Who hasn't finished? Hurry up!' Then to one child: 'Everyone has finished except you!' In short, individual progress was measured entirely in relation to the class norm *as if* this was a selection criterion, rather than by criteria dedicated to the individual. Teachers did not feel therefore that their role was to attend to the needs of the individual, since the emphasis was on passing on a body of knowledge rather than responding to the perceived needs of the child. This made the notion of 'teaching style' virtually meaningless for most teachers.

The emphasis on examination marks as criteria for success was demonstrated in that children who scored the highest examination and test results were selected both to answer more questions than others and to take responsibilities in other areas, such as helping to discipline the rest of the class and to run errands. Children in each class who

achieved the top three examination averages were awarded a much envied badge with 'honour' written on it. These were presented at a public speech day. Results of all children from grade 2 upwards were displayed publicly within school. In addition, the best results in Alexandria for the grade 5 *qabuul* examination were circulated to all the schools. EGC as a school therefore competed with other schools, vying for top position in the zone (local area). In 1996, grade 5 pupils from EGC took five of the best overall marks in all Alexandria. This appeared to give great pride to children, parents and the school *as if* it held selective value.

### ***The role of objectivity in assessment for selection***

It was believed that by sheer hard work and the Will of God, any child could succeed in the examinations, since they were considered undistorted by emotion or personal bias. The belief in this lack of bias contributed to the almost total reliance on examination results as the basis for competition and selection. The reliance was further encouraged by the large numbers of children and the government's demand for detailed monitoring information (see below).

The objectivity of teachers or inspectors who wrote the examinations remained largely unexamined, so long as rigid regulations regarding the secrecy of examination papers and their marking were adhered to. That is, the examination system was objective, while the instruments or means of assessment were only assumed to be so. Emphasis on objectivity contributed to a need for close accountability.

No-one other than the First Teacher herself saw school based examination questions until the day of the examination. The examination was therefore objective in that she did not favour individual children; but its objectivity and validity and her skill in developing a valid test were assumed only because she was a First Teacher. The same regulations applied even when the First Teacher had no previous experience of examination development, as was the case with the First Teacher of English for grade 4 in 1996. After the examination, some questions of validity were raised regarding the paper she had written. By then, however, marking was completed, according to the model answers she herself had provided, and any changes would have caused parental suspicion and fear.

The process of administration was outwardly fair, allowing little room for favouritism to individual children in the form of helping them answer questions in the examination room itself. All examinations for grades 2 to 5 were administered in *legnas* or special examination rooms in which the desks were well spaced to prevent copying, unlike during school lessons. Each child had a card with a desk number on it; on the day of the examination, children had to find their desk with its number and sit the examination surrounded by children they might not know and supervised by two teachers they might not know. In the certificate examinations of grades 3 and 5, the teachers who invigilated came from a different school and no EGC staff were allowed near the building where these examinations were held.

Despite such strict security rules, it was not uncommon for the two invigilating teachers to stand and chat rather than to keep a close guard on the children. Some pupils said that the teachers who came from outside to invigilate the grade 3 and grade 5 certificate examinations actually helped them when they could not answer a question.

Coded identity numbers were allocated to examination papers and pupil names removed. Teachers marked the papers, being fed one batch at a time by the First Teacher. The theory was that no teacher could tell whose paper she was marking and so could not show favouritism. After each paper was marked and double checked, it was returned to the safe in the Control office where marks were once again checked. For grades 3 and 5, marking happened at the zone headquarters, so teachers from a variety of English medium National Institutes marked the EGC papers.

Children's papers were therefore anonymous and marks thoroughly checked to protect objectivity. Yet, mark allocation and marking criteria were less objective. Firstly, no items were piloted before being used: the way items were presented and the variety of acceptable answers were not decided by objective criteria. In some cases, adaptation was allowed to the Model Answer Key in light of children's actual responses, but more usually the First Teacher who wrote the paper had unilateral and predetermined control over marking criteria. It was not unusual that pupils showing alternative, yet still mathematically sound, answers to a mathematics item, were marked wrong if the answers did not appear on the Model Answer Key. Secondly, although every mark had

to be checked by a second teacher-marker, in most cases it was the adding up of marks that was checked, rather than criteria used.

The whole assessment system was based on examination marks and for this reason objectivity was emphasised. And yet, between prescribed rules of practice, random or idiosyncratic practices still flourished; and these had the potential to make a mockery of the very rules implemented to protect against them.

### *Selection and social elitism in the case study school*

There were three senses in which all assessment at EGC played a selective role: EGC itself was a school for already selected children; they were educated using an elitist language; and their parents could usually afford the luxury of private lessons to put children at an advantage over other Egyptians. In these senses the whole assessment system was biased towards a select few, despite some attempts to disguise this elitism in a show of objectivity.

Initial entry to the EGC was theoretically objective, yet in practice selected only the more socially elite children for entry. The school was obliged by the Ministry to accept that, to apply to EGC, children had to be at least four years old on the first day of October of the school year in question, and not more than six years old for kindergarten 1. In the 1990s, there were usually between 700 and 1,000 applicants for about 150 places at EGC. The oldest pupils, but also children of teachers and those with siblings at the school, were offered these places. The older children had sometimes been to other schools first so parents who could afford to keep their children in kindergarten for more than two years stood a better chance of having their daughters accepted to EGC.

Additionally, although its fees were low, fees at EGC were higher than the costs of government schools and only parents working in the private or the professional sector could afford them. Further, teachers at EGC received automatic entry for their daughters, which meant that these parents were all university graduates. In these ways, EGC selected a privileged population of children whose parents had more money and were better educated than the majority in Egypt. The full details on the parents interviewed at EGC, and the effects social elitism had on their assessment orientation, are documented in Chapter 7.

Because of the social class of children who entered EGC, school-written examinations were harder at EGC than in government schools; and at EGC, examination success was complicated by being partly through the English medium. Parents' dedicated support was essential to keep children up to the elite standards set by the school. They had either to be rich enough to pay fees to private tutors or competent and available to tutor their own children consistently. Children of uneducated parents with only enough money to pay school fees stood little chance of achieving the high marks associated with success. In this way, social elitism was maintained, using tests and examination results as a gauge of how much support children were given in addressing an excessively packed curriculum.

The essential role played by parental support in the child's school achievement was recognised by teachers in class. Some teachers admitted providing clear, detailed examples of work so that parents could tutor their children using the same method as the teacher. Others openly appealed to children to ask parents to practise a skill with them. Occasionally children would cease listening in class, and teachers were accused by parents of not bothering to explain well in class, knowing that the material would be explained at home.

Another means whereby social elitism was maintained was the use of English in learning and assessing. Although Arabic was the language spoken the majority of the time and the profile of Arabic was high, a good command of English was associated with high social status. Teachers at EGC who did not speak English, who taught religion or sport, for example, had a separate staffroom from the English speaking teachers and were generally considered to have less status.

Parents who had weak or no English may have chosen to send their children to EGC in order that the children acquire this prestigious language. Their children, however, missed out on the extra support that a parent could give in English. Such children struggled more than others to understand the subjects taught in English; and so achieved less well in examinations than their peers. In addition, their parents might feel hesitant to communicate with their child's English speaking teachers. It was noticeable that parents who did not speak English were reluctant to come into school to be interviewed, while this was not the case with those whose English was strong. In this

sense, then, while examination marks were considered objective bases for competition, through use of English, good marks in the examinations were partly dependent on belonging to an English speaking elite.

To a certain extent money, however, could buy social status via examination results. If parents could afford both the school fees and private lessons in any subjects where their child could not keep up with the class norm, such as English language, then the child might perform acceptably in the examinations. Private tuition was the backbone and the bugbear of the assessment system. Private lessons were considered the antithesis to objectivity. To obliterate favouritism from private lessons, whereby private tutors gave higher marks to pupils who paid more tuition fees, the Ministry had decreed that Teacher Assessment no longer count in the accumulation of marks; no teacher of a class could write the same class's examination; and in addition, no teacher of a child could tutor that same child privately. Every teacher had to sign agreement with these regulations.

Since private lessons were so essential and yet officially discouraged, both parents and teachers were unwilling to discuss them openly. In general, though, it seems fair to say that parents tried to avoid entering their children for private lessons if they could. At primary level, lessons were considered time consuming, expensive and unnecessary if the parents themselves could help their own children to study. All parents saw themselves as playing some role in the child's examination preparation in terms of study and it was only when they could not fulfil this role that private lessons became a resort.

Many teachers, however, did give some private lessons and parents preferred the class teacher to give their child lessons, even though by law this was forbidden. One teacher [AEL, interview notes], who admitted to being very tired near the end of the year, said she taught groups of four or five private pupils most evenings after work, charging about LE20 (£4) for two hours per child. The children came to her house for tuition. Some came because they were struggling to keep up in class; others because they wanted to be top of the class.

Details about the actual uptake of private lessons are reported in Chapters 6 to 8. In this section, it suffices to say that private lessons were a means of enhancing school performance and of perpetuating social elitism because of their cost. They also fuelled

government attempts to prevent favouritism. But while examination results were the overriding indicators of success, those in elite positions capable of fighting for that success would find ways around government prohibitions.

### *The negative role of assessment for selection in learning and teaching*

In Chapter 2, some negative effects of high stakes examinations were noted. Many of these have already been mentioned with reference to the selection-based competitive assessment of EGC. They fell into two interrelated categories with regard to educational effects: fragmentation and restriction of curriculum due to emphasis on test results; and emphasis on rote learning rather than learning for interest or for relevance to life outside school, which could mean that pupils came to lack intrinsic motivation. Both were liable to be accompanied by anxiety caused by the pressure of competition and an increased likelihood that those who did not succeed would feel socially as well as academically inferior. These two negative educational aspects are addressed in detail in the following sections, with particular reference to mathematics which was the subject on which observation was based.

#### *Fragmentation and restriction of curriculum*

Since the aim of teaching was to cover the government curriculum so that children could pass a written assessment of it, it was not always perceived necessary to relate the syllabus content to the concepts themselves on which the syllabus hinged; nor was it perceived necessary to relate early content in the syllabus to content appearing later, unless one depended on the other. Material tested in January was not tested again in May. This tendency to teach closely to the syllabus, unit by unit, was exacerbated by the time restraints on teachers to complete the syllabus in time for each set of assessments. The syllabus was divided into two equal parts, one for each set of examinations in January and then May. Yet in 1996 the second half of the syllabus had to be taught in considerably less time than the first since national and Islamic festivals took more than a week out of April and this had been overlooked by those who distributed the curriculum. This overload discouraged deviation from the page of the text, even when time spent consolidating basic concepts might have saved time in the long run.

Subjects were kept strictly apart, especially the English medium subjects from the Arabic medium subjects. In addition, non examination subjects were voluntarily sacrificed if an examination subject needed extra teaching time. The possibility was little entertained that subject content was coherent within itself across the years or that subjects, including practical subjects, could relate to and support each other in the child's learning. An emphasis on operation rather than context was particularly noticeable in mathematics when children were solving 'problems'. The use of English as a second language distanced children from meaning in problems; but this distance was increased by a tendency to attend mainly to the mathematical operation demanded by the problem, rather than the problem itself and why that operation was appropriate. Where the problem involved dividing up sweets, the sweets were often forgotten when the answer was presented.

In some senses, learning test taking skills trained children to work out which operation was called for to produce an answer. Teachers working through revision questions often said to the class, 'What do they want?' as if the test writer was trying to catch them out. Its limitation was that children did not always understand why a particular transformation needed a particular operation. For example, when converting grams to kilograms, children were given the formula 'from big to small: divide' without being shown that division by necessity involved working 'from big to small'. In one geometry class, a child asked whether an equilateral triangle was not also an acute triangle; rather than receiving praise for making this connection, she was told that *if asked* about sides, she should call it equilateral but *if asked* about angles, she should call it acute.

### *Rote memorisation*

Classes consisted to a large extent of information being imparted from teacher to pupils; they sat in rows of tightly packed desks and received the information for learning. Rough working had to be done on a separate piece of paper or to be erased so that children studied a finished model. This was why parents or private tutors had to spend time each evening revising and repeating what the child had been transmitted in class. Test taking strategies, as well as the material in the book, had to be rote memorised for the tests and examinations.



The use of rote memorisation was well demonstrated in a grade 4 class learning long division [SMA, observation]. They learned to chant the steps of the operation in unison as a class. They learned where to write each number. They learned where to do their working out. They learned how to check their answers. It was all written in their copybooks and they studied these at home with their parents. By the time examinations were due, many of them could correctly answer long division sums involving five digit numbers. But, during observation, it was never obvious that any of them understood what the five digit number represented, nor what division meant, nor why the answer was as large as it was. Children learned off by heart that 'division is the opposite of multiplication'.

There were teachers who tried to emphasise children's problem solving skills. One had the motto, 'Read, then think, then answer' [SMY, observation]. She encouraged children to approach items from different angles. But she too was under pressure to give the children in her classes the best opportunities to excel in tests. She too was bound by law to teach the whole of one particular unit in a particular month. So although her pupils learned less by rote and in a less fragmented manner than some, her objective was still that children achieve the highest marks in tests and examinations.

## **5.4 The role of assessment for accountability and system improvement in the case study school**

### **5.4.1 Assessment for accountability or improvement in the case study school**

Indirectly, measures for accountability enable improvement, in that when a school is found not to be meeting required standards, intervention takes place to improve standards up to the required level. By *checking up* on the whole system of education, the controlling authority gains guidance on the *building up* of the system (see Chapter 2). The emphasis at EGC, however, was on checking up that central policy was continually being implemented by its mediators rather than on building up educational productivity or evaluating reasons for lack of improvement. This was because national reforms, using assessment, targeted government schools, including those in deprived rural areas. In those schools, accountability to the government was directly related to improvement in the schools. In the case study school, however, accountability to government in some ways limited educational programmes rather than improved them. Therefore, because of the government's need to maintain conformity to the central power of government, the accountability which enabled improvement in some schools, merely enforced conformity in others like the EGC.

The MoE believed that making the curriculum of a high standard, the national standard of all schools would be raised and the proof of this would be good results from examinations which assessed this curriculum. It was a model of minimum competency testing and the testing of basic skills. The Ministry was eager that no primary schools slipped through the net of accountability. If the Ministry had inspectors in all schools and monitored the examination results of all schools, they could be sure that no school was below a minimum standard. They also assumed that certain basic skills needed to be accounted for before children were allowed to move on to the application of skills. For both these reasons, the grade 3 and grade 5 external examinations were introduced in all primary schools, alongside the system of inspection. Improvement in this sense meant quantitative improvement of the national average attainment in primary schools.

Another sense in which the grade 3 and grade 5 examinations targeted improvement of a more qualitative nature was that the new curriculum which they assessed emphasised more than before, higher order skills and their application, deep-level learning and the

construction of knowledge in the child's mind (see Chapter 4). This emphasis was partially an attempt to match the emphases in curricula of other influential countries, and NCEEE personnel who promoted this emphasis, had been influenced by British and American educational consultants. There appeared to be a tension, however, between this emphasis and the perceived need for the curriculum to contain as much quantity of information as other influential countries, or perhaps more information, which was presented to Egyptian children at the earliest possible age. For example, English as a foreign language had originally been left out of the government primary curriculum so that pupils could concentrate on their native language, Arabic. In 1996, however, it was nationally introduced at grade 4.

Tensions among these three forms of national improvement were obvious in the case study school, where teachers perceived reforms from a different perspective:

1. Teachers at EGC felt the standards which the grade 3 examination promoted were beneath their own because they were more appropriate for underprivileged schools. Teachers did not relate to the need to assess basic skills before applied skills, since the curriculum had never been so carefully graded for progression and the new curriculum was only partially progressive.
2. Teachers at EGC were undisturbed by the new emphasis on higher order skills, perhaps precisely because this had been combined with a national drive in achieving minimum competency in basic skills: higher order skills were more difficult to prescribe and assess so local inspectors who wrote the external examinations chose not to emphasise these. This situation had been worsened by the NCEEE's failure to bring all 26 governorates' educationalists together in a common approach to assessment. During inspection, many governorate inspectors chose not to encourage this new, more complex, way of thinking.

Teachers therefore did not have to stress higher order skills in their teaching. Nor did they have the opportunity to model their examination questions on those in external examinations, when they came to write their own school examinations. Headteachers could have shared with First Teachers the NCEEE's *taqwim* or guides issued with the curriculum which laid out assessment objectives and gave example items (see Chapter 4), including items demanding higher order thinking, but this rarely happened. Had

teachers received any of these forms of training, it is possible they would have focused more on deep-level learning and less on rote memory of the curriculum. Gradually, their beliefs about learning might have changed.

3. Teachers at EGC related best to the concept of improvement as the transmission of increasingly sophisticated material and so long as they were still achieving excellent examination results based on a difficult curriculum, they felt they were making improvement. They were aware that their own school examinations were harder than external examinations, so they could relax in the knowledge that their pupils were improving in line with the improved curriculum.

Teachers sensed that central authorities' provision for improvement did not apply to them. They felt they were doing better than the national average and therefore what the government prescribed was not relevant. This reflected a trend in Egyptian education whereby minimal research was carried out at school level. Many teachers complained that no-one ever asked what they had found out in the classroom. This was indeed the case with the 'improved' curriculum which had never been trialled in classrooms, nor had its developers been into classrooms [USAid curriculum trainer, interview notes]. While the NCEEE aimed at monitoring differences across provinces, as described in Chapter 4, evaluations of reasons for failure or lack of improvement were not a common practice.

#### **5.4.2 The means of assessment for accountability**

There were two main means of assessing for accountability in the case study school: inspection by Ministry inspectors and their internal supervising agents; and submission of examination results. The inspection was often a monitoring of the process by which the results were produced.

##### ***Assessment for accountability through examination results***

A close account was kept of each child's examination marks for all subjects and in all grades. Children were accountable via examination marks to parents, teachers and ultimately, the local branch of the MoE. As described below, these results were so important that strict monitoring was needed to ensure that they provided a standard

measure of success. This monitoring was carried out by Ministry inspectors and their internal agents.

#### *Accountability and examination procedures*

Because examinations were one vehicle by which account was kept of the school and of teachers and children, and because of their importance as objective selection mechanisms, procedures which made the school accountable were extremely strict during the examinations. These procedures are discussed below with reference to individual children and to teachers separately.

#### *Accountability of individual children*

Within examination halls or *legnas*, each pupil was accounted for with severe attention to detail. As the school year began in September, the students' affairs department in charge of examination administration, sent to the Ministry a list of the names of all children eligible for examinations in the forthcoming school year. In December and in April, as the examinations approached, the Control Department sent another list, called 'list 12', of all children in the school intending to sit the examinations. The Ministry checked this against the September list and signed approval if the lists concurred. From list 12, once approved, seat numbers could be designated by the Control department. Using the Arabic alphabetic order, a seat number was assigned to each child in each primary grade.

All grade 2 and 4 children received a final report informing parents about examination results. For grades 3 and 5, final results were released by the zone when all papers in the zone had been completed, but no reports were written for these grades within the school. The school was also accountable to parents in that any parent who felt that a mistake had been made in the marking of the child's paper could complain to the school. For a considerable sum of money the parents might then be permitted to check on the marking of the paper in question. If their complaint was justified, their fee was returned. If their complaint was misjudged, the school kept the fee. Alternatively, if parents won the sympathy of the relevant inspector in the zone, he could check up on the papers himself, but this was a lengthy and arduous process.

### *Accountability of teachers*

During the examinations, a strict account was kept of invigilators so that favouritism by or to individuals was precluded. This was essential so that examination results be seen as fair. One senior and one junior teacher was assigned to each *legna*. At the end, the senior teacher signed her name as she returned her set of examination papers to the headteacher. The papers were then taken to the Control room, usually under the inspection of the government inspector. When marking was about to begin, examination papers were collected from Control and the First Teacher signed for the number of papers she had taken. She also took also copies of the Model Answer Key for the paper and an empty list with spaces for the names and signatures of each teacher marking each question.

Signing one's initials was a frequent and essential aspect of school accountability procedures since it was felt that, through a signature, the teacher made herself accountable. It was paramount during examination marking where teachers' agreement with measures protecting objectivity were most vital. Having marked one section on a paper, the marker signed her initials. Her partner then checked the section and signed her initials too. The first marker, however, was bound to use red biro, while the checker had to use blue. When the marks for each question had been added up, the teacher who did the adding signed in red, and her totalling was checked by another who signed her agreement in blue.

Once marked and totalled, each batch of examination papers was returned to Control by the subject First Teacher. The Control staff then checked the adding of all totals and summonsed the First Teacher, not the individual marker, to discuss any queries. At this point, accountability shifted away from the individual teacher and onto the internal supervisor again. Meanwhile, the First Teacher herself was closely monitored in her actions by the Ministry inspector. Once children's names were remarried to their seat numbers, pupils' names and marks could be entered on the pre prepared list. This list was submitted to the Ministry as a final account of the school's performance.

### *Inspection by Ministry inspectors*

When asked how often inspectors came into school, the headmistress replied exasperatedly that they came every day [ID, interview notes]. There were inspectors for every subject and every phase. Often they came unannounced and irregularly. Especially when an inspector was new, he came more frequently to get to know the teachers. A mathematics teacher, on the other hand, who had been in the school for a year and a half said she had never even spoken to the mathematics inspector [HML, interview notes].

Inspectors were obliged to give some training to new teachers and to explain to heads of departments about government innovations [First inspector, interview notes], but primarily their role was to monitor that schools were in line with the national curriculum prescriptions. Each inspector looked after about four schools, including completely private schools. When they came, inspectors had absolute power as Ministry employees, even when dealing with teachers who were more experienced than they were themselves. The headmistress of the junior school, QM, who had spent her whole life teaching mathematics at EGC had to show the mathematics examination she wrote for grade 4 to the First inspector; he then asked her to change its structure to be in accordance with others in the zone [QM, interview notes]. A First Teacher in grades 2 and 3 had to show the inspector daily lesson plans and was asked by the inspector to increase the contents of each since the Ministry had recently stressed the importance of lesson plans for all teachers [HMA, interview notes]. Some inspectors were more vigilant in particular duties than others: the primary science inspector, for example, was noted for his insistence on seeing children's copybooks every time he came so that he could check they had covered enough material and in the approved form [ASM, interview notes]. Although it could be considered intrusive, for less experienced teachers and especially in the field of test development, this detailed monitoring could be a useful way to provide training in a system where very few teachers had initial training (see Chapter 8 regarding teachers' backgrounds).

It was not surprising, though, that teachers in the case study school regarded the present research project which focused on their school and their assessment procedures from a considerable distance. They genuinely suspected that the headmistress wished to

monitor their behaviour through a foreign agent so that she could weed out subversive or inadequate individuals. In this system of social and political control through assessment, the teachers and the school had neither a voice in policy formation and implementation; nor did they stand to gain by complying with prescriptions except in receiving no penalties. These circumstances also explained why sometimes teachers made minimum effort to implement the government's policy, only enough to appease those who kept the account. Monitoring and accountability of this nature and the behaviour it encouraged was common to many departments of Egyptian life other than education, most noticeably in the vast government job sector. Incentives for maintaining the status quo currently outweighed any benefits of striving for a different kind of improvement; and anyway, it was not an Egyptian characteristic to believe that, under the Will of God and the socialist state, the individual could really make any difference.

#### *Financial accountability*

One striking feature of the strict monitoring for accountability in the case study school was that it was not a financial accountability. Financially, the school was only monitored to check that it was not charging higher fees or paying higher salaries than prescribed by the government. The emphasis was on financial uniformity across schools, not on accounting for government inputs since EGC received no government funding. Assessment through inspection accounted rather for the relationship between government prescription and its implementation.

#### *Accountability and curriculum distribution*

Curriculum distribution was monitored through the inspectors. Apart from teachers having to test each unit of each subject text book each month in order to prepare children for monthly testing, a check was also kept by government inspectors that the centrally prescribed number of lessons was being assigned to each subject. These two prescriptions and the monitoring of their implementation, allowed teachers minimal exercise of professional discretion as to when and how to teach particular topics. One teacher, for example, felt her grade 4 class were not ready to learn long division but she was obliged to teach it in the prescribed month in accordance with the central scheme;



the result was that she taught by rote to speed up and facilitate the process so that she could meet the assessment deadline [SMA, interview notes].

In theory the allocations shown in Table 8 of Appendix B were adhered to uniformly across Egypt in order that a desirable balance of subjects be taught which did not only emphasise written subjects. This did not apply to high level English and French which were voluntarily added to the national quota by EGC. In practice, accountability through examination results for children and teachers took on greater weight than accountability through inspection so that when the inspector was not in school, practical subjects were often sacrificed to academic examination subjects. The timetable was too packed to allow a balanced emphasis on both; and examination results always had priority.

### ***Monitoring for accountability through internal inspection***

First Teachers were those with longest experience who were appointed by the school to supervise less experienced teachers within one subject and for one age range on behalf of the government inspector. Every First Teacher was supposed to inspect the teachers within her supervision once a month by sitting in on their classes. The government had issued assessment criteria by which First Teachers could assess the teachers they supervised.

In addition to this form of monitoring, it was usual for the First Teacher to meet together with her team of teachers about once a month to inspect their preparation copybooks and to look at children's exercise books; this way she could ensure that all subject teachers were covering all the required curriculum in parallel and in the required way. These checks were then approved by the subject inspector as well as the head of department and the headmistress. They were mainly an indication of having conformed to government requirements although they were effective in preventing teachers from missing parts of the syllabus or lesson plans and in this sense, could be seen as instruments of improvement, too. They fed into an annual report for each teacher which was later sent to the local Ministry.

Every teacher was also obliged by law to calculate 'statistics' each month after the monthly test, which consisted of lists showing the percentage of girls achieving 90 per cent and above, 80 per cent and above, and so on. The First Teacher kept track of

every teacher's statistics and these statistics were also taken into account when the teacher's annual report was prepared. In government schools, a good annual report led to the possibility of becoming a headmistress or inspector, but reports had no real leverage in EGC since most teachers stayed there all their working lives and promotion depended on older teachers retiring. In terms of benefits for the teachers, therefore, they had little effect except to build up a reputation which might indirectly lead to privilege.

## **5.5 The role of assessment for the improvement of the individual's learning in the case study school**

Assessment for certification and selection in the case study school focused on individual children's test results as predictors of future selection. Assessment for accountability and system improvement focused on the behaviour of teachers and the school as a whole, in relation to national prescription. Assessment for the improvement of the individual's learning, however, was assessment within the process through which children achieved their results.

Assessment for the improvement of individual learning is defined in Chapter 2 as assessment whose purpose is improved learning within the classroom or learning site, regardless of requirements by local or national authorities for the submission of assessment results. Its purpose is formative, that is, the teacher uses assessments which will provide her with information most likely to help each child or a group of children learn more effectively. Although, in the context of EGC, the teacher's goal may have been to improve examination results that would subsequently be submitted to the Ministry, she did not submit the results of the assessments she used for formative purposes directly, since their purpose was to be fed back into her teaching (and perhaps to parents). Assessment for the improvement of the learning of the individual means that the individual child *or class* is the beneficiary of assessment rather than an outside authority.

In Chapter 2, it is stressed that the specific means of assessment for learning depended on how the word 'learning' was understood, which in turn depended on how 'knowledge' was perceived. These understandings could be determined by tradition and culture, empirical evidence, circumstance, personal belief or a mixture of these.

### **5.5.1 Perspectives on learning in the case study school**

Egyptian tradition, Islamic, Napoleonic and from the 1952 revolution, tended to view learning as not necessarily an individual pursuit. Equality and national unity depended on the indiscriminate transmission of a body of centrally prescribed knowledge. Materials to supplement the government text book for the weakest or strongest pupils were prohibited by the Ministry because individual differences were not officially recognised. The improved learning of the individual therefore tended to be perceived in

the case study school as the improved learning of a whole class of individuals, although some methods were observed whereby a child was individually assessed for formative purposes (see below).

Interviews with pupils, parents and teachers (Chapters 6, 7 and 8), however, showed that a lot of learning and teaching took place outside the classroom as a compliment to whole class teaching. Parents and private tutors gave individual or small group teaching and assessment, precisely because this could not be or was not provided in the case study school classroom. This section describes assessment for formative purposes in the case study classroom; but the accompanying teaching and learning outside the classroom which the interview data described, should be kept in mind.

In either case, the transmission model of learning was assumed, whereby a hierarchy of basic skills were passed from teacher to pupil; and initially, from government to teacher. The vocabulary used for studying reflected the common understanding of 'learning'. The verb relating to learning for tests or examinations was *tazaakir*, translated as 'memorise' or 'learn by heart'. The completion of homework was described as doing *wagab*, or 'duty'. The understanding was summed up by the head of the junior school who wondered why anyone would want to observe lessons at EGC. When told that the children's responses were of interest, she replied: 'They have no responses. They study what they are given.'

Specific information was prescribed for each grade of the system, and had been fixed by professors who wrote the curriculum in Cairo [USAid curriculum trainer, interview notes]. There was little sense of progression, or interdependence of knowledge between learning in one grade and learning in the next grade up (see Chapters 7 and 8). There was little sense that context was related to learning, since learning was seen as independent of the learner and her context. There was little sense that there could be more than one right answer to a test question, or that pupils could negotiate an answer.

Assessment for formative purposes in the classroom, therefore involved the teacher in finding out whether or not the learner had mastered one skill or fact, so that the learner was ready to move on to the next skill or fact in the hierarchy of the basics. The teacher needed to bridge the gap between what the children had already mastered and

what they should have mastered in order to be able to succeed in the next examination. Assessment in the case study school was rarely for the purposes of a detailed diagnosis.

### **5.5.2 Specific assessment tasks given in class for formative purposes**

#### ***Oral quizzing***

Teachers in the case study school quizzed the whole class, orally, often at the beginning of a lesson, to find out what they had remembered from a previous lesson, and whether they were ready to learn some new information. Usually, they asked closed questions and all the children chanted the answers. If they did not all respond, the teacher knew she needed to re-emphasise a particular point.

Sometimes teachers directed questions to individuals, and asked all the other children to say nothing, since they wanted to find out whether a particular individual was keeping abreast of the learning. One means for doing this was the cloze procedure approach, whereby the teacher left a gap in her sentence. For example, HML, said to the whole class, 'When I read the protractor, I start reading it from --?' Another technique teachers used for oral quizzing, was to phrase the same question in several different ways, to make quite sure that children had thoroughly grasped a concept.

At the end of a class, it was not uncommon for the teacher to ask an open question, to find out who still did not understand the lesson. Pupils appeared able to respond honestly to this question, so that the teacher was able to see who needed particular attention next time. Teachers did not seem to take down the names of these pupils.

#### ***Blackboard questions***

A method of assessing for formative purposes which was popular with pupils, was for the teacher to write questions on the blackboard. Usually, the pupils would attempt to answer the questions in their classwork copybooks first of all; then the teacher would invite individual children to solve a question each, in front of the other children. This way, she could select children she wanted formative information about, and watch how they carried out a process. If a child made a mistake, either it was corrected very quickly by the teacher, so that other children would not copy down a wrong answer; or the teacher invited a more competent child to answer the question. There were no examples of teachers taking notes about how children performed at the blackboard.

### ***Whole class recitation***

The teacher would ask children to recite a piece of text or a phrase. Sometimes the phrase children recited consisted of instructions for carrying out a process. For example, in SMA's lesson on long multiplication, the children chanted each stage of the operational process. SMA listened to check that enough children were responding and that their response was correct.

### ***Classwork and marking or supervising***

Teachers set work to be done in class. Sometimes they took in the copybooks at the end of the lesson, which meant that all children had to work at the same pace. Other times, they marked the work in the lesson: children queued up at their desk and teachers pointed out errors immediately so that they could go straight back to their desks and correct them. Sometimes teachers recorded the marks they gave in class.

HML wandered around the classroom as children worked, and tried to see how well children were doing so that she could intervene immediately. The tightly packed 47 desks in the class made it impossible for her to see all the children's books, however. For this reason, they used a system of rotation, so that the front row of children changed each week.

### ***Homework and marking***

Teachers set homework and usually they marked it in the staffroom after the class. Sometimes they went through each question on the blackboard and each child marked her own. Homework was important so that the teacher could see how well the child performed without her help, and it gave the pupil practice in a particular topic. This was a less effective means of gathering formative information, however, since many pupils worked through their homework with their parents or private tutors. Children were seen, however, to look at the questions they had answered wrongly in their marked work, after they had looked at the mark they received. Many of them would take the work home and their parents and tutors would help them correct it.

Teachers noted down homework marks in their mark books so that they could follow a child's progress. Some teachers made notes when a lot of children made a particular error, and made a point of reteaching that topic next lesson. Because of the large

classes, teachers could not mark all the work children completed. SMY set work from the children's mathematics workbooks quite frequently, then she would check that children had done the work by wandering around the class, only checking answers in a quick way. She was one teacher who grew angry with the children when they were absent, because it meant they missed a lesson from a unit of the text book; then she would sit with the children when they returned to work and give them individual attention. This was rare, however, among the teachers.

### *Weekly tests*

Most teachers gave weekly tests, which they warned children of, in advance. These assessed a small topic, covered in the past week. Teachers used these to check up that no-one was falling behind in the month's unit, as well as to motivate children to study and consolidate material in children's minds. Sometimes they would go through the answers orally in the next lesson, but at other times they checked and simply gave each child a mark for the test, which they recorded in their mark books.

### *Monthly tests*

Results on the monthly tests themselves, reported individually to parents in the report booklet, were a more weighty and standardised check on children's progress. On the basis of these, teachers could feel comfortable that their classes had reached the level required to pass the bi-annual examinations. These tests were set by the First Teacher and all parallel classes in a grade took the same test. These tests, unlike the bi-annual examinations, took place in children's usual classrooms and were marked by their usual teachers.

These more formal tests were formative for the individual, in that they enabled the pupil herself to see that she needed to go back and study a topic more, if she did badly. The teacher would motivate her to go and study more by reprimanding her and giving her low marks, although the teacher would not often point out to the child what her specific weak or strong areas were, nor note these down for herself. It was assumed that a poor mark reflected the need to study more, rather than a conceptual problem or a problem with how a task was presented in a lesson or a test. If a child was doing badly, the teacher would direct the pupil to her parents or tutors who could give her individual coaching.

### **5.5.3 Assessment as motivation for learning**

Since 'learning' was little distinguished from 'memorising', unless a child enjoyed the challenge of memorising, task interest or intrinsic motivation were not to be expected. Children did not expect to enjoy the content of their study; they worked hard to gain fulfilment on achieving the highest place in the class competition for marks, and were motivated by being with colleagues (see Chapter 6). All children at EGC were capable of performing above the national norm and none had actually selective barriers until grade 11. Their motivation was therefore founded in a desire to do at least as well as expected by parents and teachers, better in tests than others, and of course, no-one wanted to repeat the examinations or the grade in the following September. The evidence showed that this motivation inspired intensive and extensive hard work by even the youngest pupils, including those who found themselves rejected to the bottom of the mark list.

The motivation children sensed as a result of good marks was obvious during lessons and was used as a prop by teachers. When they gave a right answer in class, children's faces lit up. One exchange was overheard to illustrate this. The grade 4 teacher, SMA, asked one girl, 'Why are you being so silly?' She replied, 'Miss, I am happy because I got the answer right!' [Observation notes]. When pupils were wrong, however, often they cried, looked embarrassed or looked afraid. In class, children were publicly praised for high marks, for example by the teacher saying, 'Very good! Clap for her!' [HMA, observation.] They were also publicly humiliated for poor marks or for making a mistake on the blackboard in which case they could be told frostily to sit down; but this could act as motivation to do better next time.

Because of the large numbers in the school and in class, children also derived motivation from any individual attention they received. Children leapt up and cried out 'Miss, miss!' in eagerness to answer any question. Children's need for attention made it easy for teachers to joke with them and pupils responded enthusiastically to any display of humour or affection. This gave them the additional extrinsic motivation to please their teacher with high marks.



## 5.6 Overview: the roles of assessment in the case study school

For historical reasons, selection was the backdrop for assessment at EGC. The very real function of the grade 11 *thanaawiya aama* examination as a selection mechanism for university entrance dominated assessment policy at primary level even though selection there had no worse implications than remaining in grade 3 or grade 5 for an extra year. It also had a certification role, although this was diminishing. In primary, policy behaved *as if* assessment had the selective function that it had at leaving certificate level. The historical circumstances whereby children were selected out of school at the end of primary also left a selection orientation until the present day. This was compounded by the current situation whereby in government schools, choice of middle school was limited by results on the grade 5 examination. In both cases, selection was vicarious, it was *as if*. The manifestation of this backdrop of assessment for selection was the competitive orientation of the school. As with true selection, its overriding focus was competition for high examination results. The large numbers of children involved at both stages made the competitive examination an obvious assessment tool.

Examination results were used as they were in truly selective situations. They were believed to indicate how well the school was performing compared to other schools in the zone; how well classes in the school were performing compared to other classes; and how well individuals were doing compared to other children in the class.

In truly selective situations such as university entrance, examination marks were supposed to indicate what skills or knowledge the pupil had mastered. This belief underpinned the central promotion of a national curriculum in which these appropriate skills and knowledge were contained. The curriculum was regarded as the hope for the future. Through it, local and then national standards would be raised and Egypt could then compete with the most industrialised nations of the world. Instead of the assessment tail wagging the curriculum dog, the curriculum determined assessment content; although it was always assumed that assessment would take the form of examinations.

In addition to examination results, the government needed another means by which to ascertain that all schools were representing a good national standard. A sophisticated

system of inspection by external and internal agents was the means to ensuring that every zone, every school and every class was covering the curriculum effectively in accordance with national targets for each grade. In addition, these agents guarded the sanctity of examination results by monitoring the means by which they were processed.

Indirectly, national standards would be raised if Egypt's young people had tertiary as well as secondary education, and if the examination results required for tertiary education were demanding enough to ensure that the universities would produce only the best informed doctors, engineers and lawyers. For this reason, the peak challenge for curriculum coverage was the *thanaawiya aama* examination which demanded the learning of a vast bank of knowledge. Since this was a nationally administered examination, norm referencing could take place at national as well as zone, school and class level for the *thanaawiya aama*.

It remains to summarise the effects selection, certification, competition, national curriculum, accountability and system improvement had on the individual's learning. Since selection, competition, national curriculum, accountability and system improvement all focused on examination results, it was not surprising that learning did too. A transmission model of surface-level learning was best accommodated within an examination dominated system. The coverage of the national curriculum was more or less equated with learning. Assessment for the improvement of individual learning therefore fed into assessment for selection and certification; its formative purpose was better learning of material relevant to examination results. By making learners and teachers of the curriculum accountable, national standards of education were to improve and indirectly, this would mean national, economic standards, and the country's international competitiveness, would improve.

# Chapter 6. The roles of assessment for pupils

## 6.1 Introduction

One question which inspired the current research was whether the roles of assessment were different for different groups of individuals. In this chapter, its roles for pupils in the case study primary school are explored. In particular, pupils' negative and positive experiences of assessment in their system are recounted, and their experiences and perceptions about central regulations for assessment are discussed. The pupils' experiences and perceptions may later be compared to the experiences of parents and teachers, and all of these compared to official policy on assessment.

In the first section of this chapter, an overview of pupils' views and feelings is given. The descriptions given in the overview are drawn from among 89 pupils' writing in the whole class context and across two classes, one grade 3 and one grade 4 class. The overview provides the broader context for the accounts given by fewer pupils in interview. These 18 children were selected from three classes, including two classes referred to in the overview. The spoken accounts make up the second section in this chapter and provide detail and depth to substantiate the information taken from pupils' writing.

In this and the subsequent two chapters, quotations are given to give the reader some first hand access to data. When a quotation summarises the views of several interviewees, it is presented within a frame. Where individual pupils are mentioned, an invented name is used (while teachers are referred to using the teacher's first initial, followed by the initial of the subject they taught, followed by a letter taken from later in their name).

### **Section I. Overview: pupils' views and feelings, as expressed in writing**

## 6.2 Overview of pupils' views and feelings

### 6.2.1 Pupils' positive feelings about their schooling

Only grade 4 children (n=44) wrote of their positive feelings regarding school, some giving more than one positive aspect of school. Sixteen children gave break time as the most enjoyable aspect of school: one child summed up the preference for break by

saying, 'I like break time very much [more] than the other lessons'. Fifteen children mentioned a friend or friends as what they enjoyed best in school; eight mentioned teachers, although a further three said they enjoyed it best in school when the teacher was absent and they had a free lesson.

Fourteen children gave a subject, or subjects, as the most enjoyable aspect of school, including English, physical training, activities and art. It might be relevant to these preferences that the class teacher [AM] taught English; other than English, the favourite subjects mentioned were non academic. One child enjoyed the library best, another the gardens and another 'the parties which we make in school'.

**Table 6.1 Overview: Pupils' favourite aspects about their schooling**

<b>Pupils' favourite aspects about their schooling</b>	<b>No. of pupils (n=44)</b>
Break time	16
Friends	15
English, PE, 'activities' or art	14
Teachers	8
Free lessons	3
Other	3]

### **6.2.2 Pupils' ambitions for after schooling**

Nearly half the children aimed to become doctors. The high prestige of this job in Egypt may have been reflected in the children's inability to give reasons for wanting to do the job, other than 'I like this job'. Most said they wanted to treat sick people (a sentence which appeared in their English text book). Only six children wanted to become a doctor to follow a parent. The predominance of girls in the class may have accounted for some of this discrepancy.

Twenty-one children wanted to become teachers; unlike the potential doctors, they expressed clear reasons for their ambition. For example, three pupils liked children; two liked their teachers; one said, 'I like to mark copybooks'; one wanted to write on the blackboard; another wanted to be like her class teacher; several wanted to help little children learn; but several others simply considered teaching to be a 'good' profession.

One of the engineers was following her father's example, but another wanted to build flats, and a third liked mathematics and drawing. The accountants both liked mathematics; one of the five policewomen wanted to '... relieve society of thieves'; and

the authoress-to-be had a big imagination so she wanted to put whatever she liked in her stories.

**Table 6.2 Overview: Pupils' ambitions for after schooling**

<b>Pupils' ambitions for after schooling</b>	<b>No. of pupils (n=89)</b>
Doctor	41
Teacher	21
Engineer	6
Police woman/officer	5
Don't know	3
Accountant	2
Pharmacist	2
[Other	9]

### **6.2.3 The purpose and use of examination results**

Many pupils gave several reasons why examination results were important to them. Promotion to the next class was the most popular response (25). This response was actually the answer to why the children wanted to pass, rather than to get good marks. A pass mark was in reality not difficult to achieve. Yet children believed high marks to be important for promotion.

Twenty-four children wanted good examination results to prove their self worth: remarks such as, 'I want to be a good girl' or 'an excellent girl' were common. Others (11) gave good or full marks as achievements in themselves.

Twenty-one children also considered good marks to be important for their more distant future. They seemed to believe that good marks in grade 3 or 4 would contribute to success in adulthood. Most commonly, the children mentioned the job they hoped for: one child explained the connection between present learning and future success in saying, 'I work hard because I want to be a teacher'. One child had the high ambition to be, not just a doctor, but 'a famous doctor' and another 'to be a good painter in a big exhibition'.

Seven children wanted good marks so that their parents would give them a present; most frequently mother was mentioned (five). Father was mentioned twice, and grandfather once: 'I want to take from my grandfather the gold pound'. Sixteen children wanted more generally to make their parents happy, knowing how much

parents valued high marks. One wrote, ‘My mother loves me and she will be happy’; another wrote, ‘When I take bad marks, I’m sad and my mum and dad are sad’.

**Table 6.3 Overview: Pupils’ views on the purpose and use of examination results**

The purpose and use of examination results	No. of pupils (n=89)
To be promoted to grade 4	25
To be a good/excellent girl	24
For success in adulthood	21
To make parents happy	16
To get good, very good or full marks	11
To receive a reward from parent	7
Don’t know	1

#### **6.2.4 Pupils’ negative feelings during the examinations**

A good proportion of children did not like the conditions in which they sat examinations, most of these being in grade 3 (12). Grade 3 took most of their exams under the supervision of the MoE, so their responses to this sentence reflected this experience during the mid year examinations in January. Five children said it was too noisy. One wrote, ‘I do not like in the exams [when] they make noise and the teachers [are] shouting. I can’t write’. Three mentioned cheating, one saying, ‘I don’t want anyone to help me’. Two pupils said they did not like the teachers who came in from the Ministry. Others complained about the desks, that they were dirty, and they had the seat attached rather than separate.

Children were unhappy about difficult questions in their examinations: ‘They are too difficult and I don’t like difficult things’, one child wrote. Mathematics was particularly unpopular in grade 3 (13) and English composition across the two grades. Nine children, all in grade 4, simply did not like having to study for examinations. One wrote, ‘I don’t play and I don’t watch TV’; another that she did not like her mother revising with her. One of the three children who liked everything about examinations was the child who consistently attained low marks.

**Table 6.4 Overview: Pupils' negative feelings during the examinations**

Pupils' negative feelings during the examinations	No. of pupils (n=89)
The questions are too difficult	27
Examination conditions are not helpful	16
Mathematics is too difficult	15
English composition is too difficult	11
I don't like having to study for exams	9
Questions including words 'not taken' in class are too difficult	3
I like everything about exams	3
Examinations are too long	2
[Other	3]

### 6.2.5 Pupils' views on how to succeed

Some pupils (in grade 4 only) had more than one view on how to succeed. Over half the children saw the way to success through studying well or working hard. Nine children stressed personal quality rather than, or as well as, effort exerted: being clever, honest, gentle, careful, or just 'good'. Only four children included success specifically in exams as a means to future success.

Three girls mentioned domestic duties as a way to future success; one wrote: 'I must take care of my children, my husband and my house'. Two children said they must listen to others: to mother, teacher and to God.

**Table 6.5 Overview: Pupils' views on how to succeed**

Pupils' views on how to succeed	No. of pupils (n=44)
Study well/work hard	31
Be clever/honest/gentle/careful/good	9
Succeed well in examinations	4
Be a loving wife	3
Listen to authority	2

### 6.3 The 24 hour diary record

All the children (n=44) got up at 6am or 7am. There was some variation as to when they left the house for school, presumably depending on whether they went by bus or by car, and how far away they lived. Only three children made reference to travelling by car, while 27, over half the class, made specific mention of travelling by bus; twelve of these described waiting for the bus.

Nearly all children said they arrived home around 4pm. Some children said they slept, usually after having lunch. Some of those who slept in the afternoon, not surprisingly, went to bed late at night, in two cases as late as two in the morning. In the whole class, 10pm was the most common time for bed.

It was common for children to watch television in the evening; only 16 children made no mention of television. Twenty children watched television at 8pm, presumably the popular series 'khalti Safia wa dar'. Ten children found time to play in their evening.

The most common number of hours studying was three hours. Fourteen children claimed to do homework or study, or both, for three hours. Eleven studied for two hours while seven studied for four. A few children referred to preparing for the next day's lessons last thing at night.

**Table 6.6 Summary of grade 4 pupils' daily schedule for 24 hours**

Activity	Pupils volunteering information (n=44)
<b>Start and end day</b>	
Get up at 6am or 7am	44
Arrive home around 4pm	44
[No comment	0]
<b>Travel</b>	
By bus	27
Wait for bus	12
By car	3
[No comment	2]
<b>Leisure activity</b>	
Sleep in afternoon	12
Play	10
Watch TV	28
Watch TV at 8pm	20
[No comment	0]
<b>Hours of study</b>	
5 hours	4
4 hours	7
3 hours	14
2 hours	11
1 hour	4
[No mention of study	4]
<b>Time for bed</b>	
9pm	4
10pm 11pm	20
11pm 12pm	11
12pm or later	6
[No comment	3]