

Statement of Originality

ALTERNATIVE ASSESSMENT IN THE TEACHING OF ENGLISH AS A  
FOREIGN LANGUAGE IN ISRAEL

by

LOIS ANN GRABIN

submitted in accordance with the requirements

for the degree of

DOCTOR OF EDUCATION

in the subject

DIDACTICS

at the

UNIVERSITY OF SOUTH AFRICA

PROMOTER: PROF MM NIEMAN

JOINT PROMOTER: PROF AJ HUGO

JANUARY 2007

### Statement of Originality

I declare that **ALTERNATIVE ASSESSMENT IN THE TEACHING OF ENGLISH AS A FOREIGN LANGUAGE IN ISRAEL** is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

L A Grabin

Student number 3426-757-3

## Abstract

There are a myriad of factors that influence the teaching and assessment of English as a foreign language in Israel. One such factor is the change in perception of how the mind collects and processes information. The transition from behaviorist theories of learning to constructivist and cognitivist theories has resulted in a new concept of learning that recognises the pivotal role of the learner in the learning process. Another instrumental factor is the creation of standards in education that are intended to raise the level of academic achievement and provide equitable educational opportunities for all learners. Both of these factors have impacted the study of English as a foreign language.

Countries around the world have assumed monumental changes in the way they approach foreign language study. Following the examples of foreign countries Israel has recently implemented a standards-based curriculum that supports alternative assessment. Since this is a relatively new concept, the principles underlying standards and assessment are unfamiliar to many EFL teachers. Novice teachers, as well as experienced teachers, are experiencing difficulty incorporating standards and alternative assessment into their lessons. In the light of the above it was clear that a need existed to investigate how English foreign language teachers relate to assessment and to examine the extent to which they implement alternative assessment in their teaching. The aim of the study is to explore teachers' understanding of alternative assessment and find out how teachers incorporate alternative assessment in their EFL classes.

Against this background, an empirical investigation was undertaken. It comprised a textual analysis of ten assessment packages and two focus group interviews. The research produced a wealth of descriptive data that illuminated the subject. The findings magnified the disparity between the theory behind alternative assessment and actual practice. It was apparent that the assessment activities which EFL teachers employed, were not in sync with the principles that were being promoted by the new curriculum for EFL teaching and assessment in Israeli schools. The study discusses the results of the investigation, suggests the implications and presents recommendations.

**Key terms:**

Assessment; Traditional assessment; Alternative assessment; Assessment for learning; Feedback; English as a foreign language; English as a second language; Standards; Language standards; Standards movement; Task; Performance-based tasks

## **TABLE OF CONTENTS**

### **1.0 CHAPTER ONE**

#### **PROBLEM STATEMENT AND DESCRIPTION OF THE RESEARCH METHODOLOGY**

1.1 INTRODUCTION	1
1.2 AWARENESS OF THE PROBLEM	5
1.3 STATEMENT OF THE PROBLEM	8
1.4 THE AIM OF THE RESEARCH	8
1.5 THE METHOD OF RESEARCH	9
1.6 THE LIMITS OF THE RESEARCH	10
1.7 CLARIFICATION OF CONCEPTS	10
1.7.1 Assessment	10
1.7.2 Traditional assessment	11
1.7.3 Alternative assessment	11
1.7.4 Assessment for learning	12
1.7.5 Feedback	12
1.7.6 English as a foreign language	12
1.7.7 English as a second language	12
1.7.8 Standards	13
1.7.9 Language standards	13
1.7.10 Standards movement	13
1.7.11 Task	14
1.7.12 Performance-based tasks	14
1.8 PLAN OF STUDY	14

### **2.0 CHAPTER TWO**

#### **FROM TRADITIONAL ASSESSMENT TO ALTERNATIVE ASSESSMENT**

2.1 INTRODUCTION	16
2.2 THEORIES UNDERLYING ASSESSMENT PROCEDURES	17
2.2.1 Behaviorism	18
2.2.2 Cognitivism	19
2.2.3 Constructivism	21

2.2.4 Multiple Intelligence	22
2.3 DEFINING ASSESSMENT	23
2.3.1 Assessment and evaluation	25
2.3.2 Testing and assessment	25
2.3.2.1 Standardised testing	27
2.3.2.2 Achievement testing	27
2.3.2.3 Aptitude testing	28
2.3.3 Types of assessment	28
2.3.3.1 Base-line assessment	28
2.3.3.2 Diagnostic assessment	29
2.3.3.3 Formative assessment	29
2.3.3.4 Summative assessment	30
2.3.3.5 Ipsative assessment	30
2.3.3.6 Norm-referenced assessment	31
2.3.3.7 Criterion-referenced assessment	31
2.3.4 Measurement and measurement instruments	32
2.3.4.1 Checklists	32
2.3.4.2 Rubrics	33
2.4 PRINCIPLES OF ASSESSMENT	34
2.4.1 Reliability	35
2.4.2 Validity	35
2.4.3 Feasibility	36
2.4.4 Transparency	37
2.5 TRADITIONAL ASSESSMENT	37
2.5.1 Types of traditional assessment	38
2.5.2 Criticism of traditional assessment	40
2.6 ALTERNATIVE ASSESSMENT	46
2.6.1 Criteria for effective alternative assessment	48
2.6.1.1 Clear instructions	48
2.6.1.2 Continuous assessment and opportunities for improvement	50
2.6.1.3 Feasibility	50
2.6.1.4 Authenticity	50
2.6.1.5 Student-centered	51
2.6.1.6 Fairness	52
2.6.1.7 Improvement of learning	53

2.6.1.8 Takes place in the classroom	54
2.6.2 Alternative assessment methods	55
2.6.2.1 Performance assessment	56
2.6.2.2 Self-assessment	61
2.6.2.3 Peer assessment and group assessment	62
2.7 ASSESSMENT IN SECOND AND FOREIGN LANGUAGE TEACHING	63
2.7.1 Benefits for language teachers	64
2.8 RELATIONSHIP BETWEEN STANDARDS AND ASSESSMENT	65
2.9 SUMMARY	65

### **3.0 CHAPTER THREE**

#### **THE IMPLEMENTATION OF STANDARDS IN EDUCATION AND ITS IMPACT ON ASSESSMENT PRACTICES**

3.1 INTRODUCTION	66
3.2 DEFINITION OF STANDARDS	68
3.3 THE RATIONALE BEHIND STANDARDS	70
3.3.1 Desire for greater academic achievement	70
3.3.2 Accountability	71
3.3.3 Guidelines for teachers	71
3.4 PRINCIPLES UNDERLYING STANDARDS	72
3.4.1 Standards must be specific	73
3.4.2 Standards should measure multiple levels	73
3.4.3 Standards must be teachable	74
3.4.4 Standards must be flexible	75
3.4.5 Number of standards	76
3.5 KINDS OF STANDARDS	76
3.6 PROBLEMS ASSOCIATED WITH STANDARDS	78
3.6.1 Unclear standards	79
3.6.2 Unrealistic standards	80
3.6.3 Teaching to the test	81
3.6.4 Limiting expectations	83
3.6.5 Too many standards	83
3.6.6 Fails to consider the needs of diverse learners	85

3.6.7 Frustrating to the student	86
3.6.8 Drain on resources	87
3.6.9 Disservice to teachers	87
3.7 ADVANTAGES OF STANDARDS	88
3.7.1 Equality of education	88
3.7.2 Common educational objectives	89
3.7.3 A tool for better teaching	89
3.7.4 Raising and maintaining professional standards	92
3.7.5 A better learning experience	92
3.7.6 The example of “world-class” standards	92
3.8 INFLUENCE OF STANDARDS ON ASSESSMENT	92
3.9 THE IMPLICATIONS OF THE MOVE TO STANDARDS-BASED EDUCATION ON A NUMBER OF COUNTRIES AROUND THE WORLD	94
3.9.1 Standards in Queensland, Australia	94
3.9.2 Standards in Canada	94
3.9.3 Standards in the England	94
3.9.4 Standards in Finland	96
3.9.5. Standards in Italy	96
3.10 STANDARDS IN ISRAEL	96
3.10.1 The New Curriculum	98
3.10.2 The effect of standards on assessment	102
3.10.3 Teaching standards	102
3.11 SUMMARY	107

#### **4.0 CHAPTER FOUR**

##### **FACTORS INFLUENCING THE DEVELOPMENT IN FOREIGN LANGUAGE TEACHING AND LEARNING IN ISRAEL**

4.1 INTRODUCTION	109
4.2 THE EFFECT OF GLOBALISATION ON ESL AND EFL LEARNING	110
4.3 DISTINCTION BETWEEN THE TERMS ENGLISH AS A SECOND LANGUAGE (ESL) AND ENGLISH AS A FOREIGN LANGUAGE (EFL)	112
4.4 HISTORY OF SECOND AND FOREIGN LANGUAGE LEARNING	114
4.4.1 Introduction	114
4.4.2 A change in focus	115
4.4.3 The effect of national standards on ESL and EFL teaching and assessment	116

4.4.4 The American Council for the Teaching of Foreign Languages (ACTFL) Guidelines	116
4.4.5 The Standards for Foreign Language Learning – Preparing for the 21 <sup>st</sup> Century	117
4.4.6 Managing the Assessment Process (The MAP)	121
4.4.7 Standards and EFL in Israel	122
4.5 FACTORS THAT INFLUENCE LANGUAGE LEARNING	125
4.5.1 Affective factor	125
4.5.2 Student attitude	126
4.5.3 Motivational factor	127
4.5.4 Level of self-confidence	128
4.5.5 Anxiety factor	129
4.5.6 Feelings of autonomy	130
4.6 THE INFLUENCE OF COGNITIVE THEORIES ON ESL AND EFL LEARNING	131
4.7 THE GOOD LANGUAGE LEARNER	132
4.8 LANGUAGE LEARNING: INDIVIDUAL LEARNING STYLES AND ASSESSMENT	133
4.9 LANGUAGE LEARNING: STRATEGIES IN ESL AND EFL	134
4.9.1 Advantages of LLS in EFL	135
4.9.2 Types of language learning strategies	136
4.10 THE ROLE OF FIRST LANGUAGE ON FOREIGN LANGUAGE STUDY	137
4.10.1 Reasons for learning language	137
4.10.2 Connection between first language and second or foreign language Learning	137
4.10.3 Negative effects of first language on second or foreign language learning	138
4.10.4 Positive effects of first language on second or foreign language learning	139
4.10.4.1 L1 provides students with a sense of security and identity	139
4.10.4.2 L1 supplies a foundation for language understanding	140
4.11 TEACHING AND ASSESSMENT IN THE EFL CLASSROOM	140
4.12 ALTERNATIVE ASSESSMENT IN EFL	142
4.12.1 Language for communication	143
4.12.2 Content-based instruction	144
4.12.3 Task-based learning	145
4.12.4. Types of authentic assessment in first and foreign language learning	146
4.12.4.1 Projects	147

4.12.4.2 Performance Tasks	148
4.12.4.3 Portfolio	150
4.12.4.4 Self-Assessment	151
4.12.4.5 Peer assessment	152
4.12.4.6 Group assessment	152
4.13 EFL TEACHING IN ISRAEL	153
4.14 ALTERNATIVE ASSESSMENT IN EFL IN ISRAEL	153
4.14.1 Assessment guidelines for the English curriculum	154
4.14.1.1 Principles of classroom assessment	154
4.14.1.2 Categories of classroom assessment	154
4.14.1.3. Standards and benchmarks in assessment	155
4.14.1.4 Performance-based teaching and assessment	158
4.14.1.5 Characteristics of a performance task	159
4.15 MATRICULATION EXAMINATION – BAGRUT	159
4.15.1 Background about Bagrut	159
4.15.2 Advantages of the new Bagrut	160
4.15.3 Dissatisfaction with Bagrut	162
4.15.4 Projects in EFL in Israel	163
4.15.4.1 Project requirements	165
4.15.4.2 Project aims	165
4.15.4.3 Stages of the project	166
4.15.4.4 Project Evaluation	167
4.15.4.5 Reflection and feedback	168
4.15.5 Performance Tasks in EFL in Israel	170
4.15.6 Portfolio in EFL in Israel	171
4.16 THE CHALLENGES OF ALTERNATIVE ASSESSMENT IN EFL IN ISRAEL	173
4.17 SUMMARY	175

## **5.0 CHAPTER FIVE**

### **RESEARCH DESIGN AND METHODOLOGY**

5.1 INTRODUCTION	177
5.2 OBJECTIVES OF THE EMPIRICAL STUDY	177
5.3 RESEARCH METHODOLOGY	178

5.3.1 Qualitative research: A theoretical basis	179
5.3.1.1 Characteristics of qualitative research	179
5.3.1.2 Advantages of qualitative research	180
5.3.1.3 Criticism of qualitative research	181
5.3.1.4 Justification for using a qualitative research design	182
5.3.2 Triangulation	183
5.4 SAMPLING	184
5.5 THE ROLE OF THE RESEARCHER	187
5.6 DATA COLLECTION	187
5.7 METHODS	188
5.7.1 Textual analysis	188
5.7.1.1 Criteria used to examine and judge the influence of standards on assessment	190
5.7.1.2 Criteria used to determine whether alternative assessment measures were incorporated	191
5.7.1.3 Procedures used to analyse content	192
5.7.2 Focus group interviews	194
5.7.2.1 Advantages of focus groups	195
5.7.2.2 Focus group participants and site of focus group interviews	196
5.7.2.3 Questions for the focus group interviews	197
5.8 FIELD NOTES	199
5.8.1 Field analysis	200
5.9 DATA PROCESSING	200
5.9.1 Segmenting	202
5.9.2 Coding	202
5.9.3 Assembling a master list	202
5.9.4 Enumeration	203
5.9.5 Intracoder reliability	203
5.9.6 Broad categories	203
5.10 SUMMARY	204

## **6.0 CHAPTER SIX**

### **RESEARCH FINDINGS AND CONCLUSIONS**

6.1 INTRODUCTION	205
------------------	-----

6.2 PROCEDURES FOR CONDUCTING THE TEXTUAL ANALYSIS	206
6.3 PROCEDURES FOR CONDUCTING FOCUS GROUP INTERVIEWS	207
6.4 PRESENTATION OF THE FINDINGS OF THE TEXTUAL ANALYSIS	209
6.4.1 The composition of the assessment packages	209
6.4.2 Analysis of the assessment packages	213
6.4.2.1 The influence of standards on assessment	213
6.4.2.2 The incorporation of alternative assessment methods	221
6.5 PRESENTATION OF THE FINDINGS OF THE FOCUS GROUP INTERVIEWS.	227
6.5.1 Analysis of the focus group interviews	227
6.5.1.1 How decisions on the specific type of assessment activities to be used are made....	227
6.5.1.2 The way in which the implementation of a standards-based curriculum influence assessment practices in EFL	229
6.5.1.3 The extent to which EFL teachers still utilise traditional assessment methods and test items?	231
6.5.1.4 The extent to which assessment practices make provision for the assessment of all the language skills	232
6.5.1.5 The classroom opportunities students have to be assessed by engaging in social interactions in authentic settings and situations	234
6.5.1.6 The assessment methods and activities that are used on a regular basis	235
6.5.1.7 The extent to which students are involved in classroom assessment	236
6.5.1.8 The extent to which assessment activities encourage the improvement work in progress	238
6.5.1.9 Your opinion on the use of criterion-referenced versus norm-referenced assessment	239
6.5.1.10 The main reasons why you might be reluctant to use alternative assessment activities	240
6.5.2 Summary of the results of the focus group interviews and the assessment packages	242
6.6 SUMMARY	244

## **7.0 CHAPTER SEVEN**

### **SUMMARY OF THE STUDY, IMPLICATIONS AND RECOMMENDATIONS OF THE EMPIRICAL STUDY**

7.1 INTRODUCTION	247
------------------	-----

7.2 SUMMARY OF THE STUDY	247
7.3 SYNOPSIS OF FINDINGS	252
7.4 RECOMMENDATIONS	255
7.5 CONTRIBUTIONS OF THE RESEARCH	258
7.6 LIMITATIONS OF THE STUDY	258
7.7 CONCLUDING REMARKS	259
BIBLIOGRAPHY	261
APPENDIX 1 Textual Analysis - Criteria used to determine whether alternative measures were incorporated	281
APPENDIX 2 Textual Analysis – Criteria used to judge the influence of standards	282
APPENDIX 3 Textual Analysis – Percentage distribution of semester grade	283
APPENDIX 4 Transcript of Focus Group Interview 1	284
APPENDIX 5 Transcript of Focus Group Interview 2	291
APPENDIX 6 Master List of Code Words: Focus Group Interviews	301
APPENDIX 7 Data Analysis of focus group interviews	302

## CHAPTER 1

### PROBLEM STATEMENT AND DESCRIPTION OF THE RESEARCH METHODOLOGY

#### 1.1 INTRODUCTION

A great deal of development has been seen in the field of foreign language instruction and assessment over the last several decades. This has been the direct result of the global spread of English over the last 40 years. Hasman (2000:2) points out that world-wide over 1.4 billion people live in a country where English has official status, and one out of five of the world's population speaks English with some level of proficiency. By 2010, he estimates, the number of people who speak English as a second language will exceed the number of native speakers, thus exercising a significant influence on the language. It is not surprising, therefore, that new approaches to language instruction have been adapted. Whereas the focus on language instruction was previously on the elements of language, grammar, vocabulary and pronunciation, the focus needed to shift to the teaching of communicative skills that could be used in authentic language settings (Bachman, 1990:27).

The standards movement is one of many movements which are credited with providing the impetus for the revolution in foreign language teaching. Standards provide students with clear benchmarks for what students should be able to accomplish, while providing them with learning situations in which they can strive to meet their educational objectives.

Standards-based reform, in general, “envisions school systems driven by agreements about what every student should know and do; guided by data from assessments which tell whether or not students have adequately learned those things; and motivated by accountability mechanisms designed to assure every student access to effective opportunities to learn” (Jamentz, 2001:3).

Countries worldwide have adopted standards in varied subject areas in an attempt to guarantee equal learning opportunities for all students. An additional aim of standards proficiency has been to provide young people with the skills and abilities that will best enable them to assume their places in a modern and ever-changing world.

The standards movement has had a considerable influence on language instruction. Language standards have become the core of foreign language instruction. As foreign language instruction changed direction, the need for different approaches to assessment became apparent. This was the case in standards-based language teaching in particular because standards indicate that language should be used in authentic settings and for communicative purposes.

Traditional forms of language testing were unsuitable for measuring the performance of communicative skills. The usual tests, such as multiple-choice, fill-in or true/false assessed the students' ability to recognise the correct answer rather than produce it (Popham, 2003:81-2). The symbiotic relationship that exists between standards and assessment, pointed to a need for alternative assessment opportunities. It seemed clear that as instruction methods changed in order to incorporate new standards principles, assessment strategies needed to be adjusted accordingly.

Cohen (1990:47) views traditional assessment strategies as being inadequate as a measurement for the depth and scope of education that a student receives. He offers several complaints to explain his dissatisfaction with traditional assessment, namely:

- Traditional assessment focuses upon products of learning, but rarely upon processes of learning.
- Traditional strategies assess limited aspects of academic learning.
- Traditional assessment strategies provide some information about certain types of learning progress in some subject areas, but largely neglect crucial areas such as, divergent and creative thinking, as well as critical thinking and problem solving.
- Traditional assessment strategies can cause irreparable lifetime damage to individual students, especially to their self-images, by providing

negative feedback.

- Many (if not most) important educational outcomes cannot be measured, weighed or counted, and in fact, assessment represents probably at best ten per cent of the total impact of learning on students.

Torrance (1995:1) faults traditional assessment with failing to stimulate students and pique their creativity. He champions a more practical, realistic and challenging approach to assessment than in the past. He criticises the “narrowing effect” that traditional tests can have on curriculum content and teaching methods. This effect, he contends, causes children to possibly know certain things without really understanding them and without being able to make generalisations from specific examples to similar problems in different contexts. Torrance (1995:2-3) argues that assessment needs to consider higher-order skills and capabilities such as problem solving, investigation and analysis. Therefore, he believes, there is a need for more ‘authentic’ or realistic tasks than traditionally used.

The result of advanced theories in language learning, has led to a greater emphasis on the acquisition and application of knowledge, rather than the accumulation of isolated facts. The cognitive revolution, according to Gardner (1987:393), began as early as the 1940’s as a response to the behaviorist approach which “avoided discussion of the brain, rejected concepts of mental representation, and averted consideration of higher-level perceptual or problem-solving processes”. The behaviorist theories have been superseded by cognitive psychology which Gage, Yekovich and Yekovich (1993:4-5) define as the scientific study of mental events. They define the cognitive psychology of school learning as the scientific study of mental events that occur in learners and teachers during the school process. They believe that in order to maximise the learning potential during this time teachers need to consider four important questions.

- What is going on in the learner’s head, for example, when he or she is learning a new principle or a new concept?
- What underlies a student’s desire to learn?
- What determines a student’s learning proficiency?
- What mental processes support routine and creative problem-solving...?

Constructivism is the second major approach to learning supported by the majority of teachers and educators (Stahl, 1996:2). A fundamental principle of constructivism is that learners formulate meaning based on their previous background and their prior knowledge and as a result different learners construct different meanings (Olsen, 1997).

The approaches mentioned above called for a new form of assessment that would present learners with the challenge to create meaning on their own. Alternative assessment became an effective way to assess how learners were able to use their skills and knowledge in many different situations. Assessment needed to guide learning and provide sufficient opportunity to practice, refine and master the tasks we want them to do (Wiggins, 1993:8).

As foreign language instruction changed direction, the need for different approaches to assessment became apparent. This was the case in standards-based language teaching in particular, because standards maintain that language should be taught for communicative use in authentic situations. Therefore, alternative assessment was important since it provided opportunities to use authentic language in real-life settings and situations.

Following the initiatives taken by numerous educational systems worldwide, the Pedagogical Secretariat of the Ministry of Education in Israel approved a standards-based curriculum for the teaching of English in July 1998. Prior to this time, the educational system in Israel had not included curricula based on standards. The standards were set after an extensive study of language learning principles that have contributed to successful language instruction in other countries around the world (Steiner, 2000:12). Since great emphasis is placed in Israel on the study of English as a foreign language, it is considered extremely important to provide quality instruction in the field of EFL. English is taught as a foreign language in schools in Israel from as early as the first grade though it is a required subject in schools from grade four through grade twelve.

While Hebrew is the official language of the country, Arabic is recognised as the language of the Arabic-speaking minority. English is considered the first foreign

language in the country. Therefore, great emphasis is placed in Israel on the study of English as a foreign language. It is optional in the 3<sup>rd</sup> and 4<sup>th</sup> grade and compulsory throughout the rest of the school system. The pre-1960 concern for literature and culture in the teaching of English has shifted to an emphasis on English as an international language of communication (Spolsky & Shohamy, 1996). In addition, Spolsky and Shohamy explain the increasing stress over the years on oral ability as a direct result of the wave of English speaking immigrants in the 1970s who entered the English teaching profession. It is considered extremely important to provide quality instruction in the field of EFL so that students will be able to communicate in English and have the language proficiency that is necessary for future study that will prepare them for a wide range of professions.

Steiner (2000:9), the Chief Inspector for English Language Education for the Ministry of Education in Israel points to two central reasons for the decision to set national standards for the curriculum in English:

1. National standards allow for pupils to have equal opportunities. They are judged by the same standards.
2. National standards clarify what pupils should know at different levels of their education.

As an English teacher with many years of teaching experience, the researcher is very interested in and affected by the monumental changes that have been taken to improve language instruction and assessment in Israel. In addition, as someone who is active in the field in Israel, the researcher has a close working connection with numerous language teachers. As a result of on-going communication with colleagues in person as well as through the frequent postings on the English Teachers Network in Israel (ETNI) site, it is apparent that many teachers experience difficulties in understanding and implementing alternative assessment.

## **1.2 AWARENESS OF THE PROBLEM**

The standards movement worldwide has had a direct influence in Israel on the

teaching of English as a foreign language. The first formal statement appertaining to the place of language learning in Israel appeared in 1995 in the document *Policy on Language Education in Israeli Schools* issued in the Ministry of Education Director-General's Circular dated June 1, 1995 (Ministry of Education, Culture and Sport, 1995) and, in a revised form, re-issued in April 1996 (Ministry of Education, Culture and Sport, 1996). The policy which took effect in September 1996, dealt with second and foreign language education as well as with mother tongue teaching. The policy justified and revised existing policies and applications of language teaching. A distinct objective was to set general goals for the instruction of English.

Whereas the teaching of English in the pre-1960 years was concerned with culture and literature, the 90s shifted the stress on English as an international language of communication. There has been greater emphasis placed on oral abilities over the years. In addition, English was recognised as the country's "first foreign language" as defined in the *Policy on Language Education* (Ministry of Education, Culture and Sport, 1995). This means that English is taught as a foreign language in all Israeli schools. Hebrew is the official language of Israel and Arabic is considered to be a second language as it is spoken by the large Arab population in the country.

The revolution in language teaching in Israel, however, must be traced to July 1998, when the Pedagogical Secretariat of the Ministry of Education, Culture and Sport approved a standards-based curriculum for the teaching of English in Israeli schools. The official draft copy of *English Curriculum: Principles and Standards for Learning English as a Foreign Language in Israeli Schools* was published in November 1998. This document outlined a number of language teaching principles that were intended to facilitate language proficiency. The curriculum is based on standards that are defined as "a cumulative body of knowledge and set of competencies" (Ministry of Education, Culture and Sport, 2001:14). The premise was adopted that in order for students to exhibit proficiency in various areas of language usage new methods of assessment were warranted. Traditional evaluation methods such as multiple-choice, short answer and true/false questions, which were able to assess what a student knew, had been considered acceptable in the past. Once the focus shifted to language proficiency and attention became focused on what students were able to do in and with the target language, paper and pencil responses were no longer adequate.

Therefore, performance-based tasks, which required the student to utilise his language skills in a true setting became appealing.

Due to the newness of this concept, EFL teachers in Israel need to have an understanding of the significance of having a national curriculum based on standards. It is necessary to make them aware of the advantages that a standards-based curriculum can have for learners. Furthermore, teachers need to grasp the implications that standards have on teaching, learning and the assessment process (Steiner, 2000:13).

In addition, as the standards movement in the last decades ushered in a different emphasis on language objectives, there is a need to develop new assessment methods to check whether the desired objectives have been reached. A problem lies in the fact that English teachers in Israel are required to assess their pupils with alternative methods of assessment, but they are not sufficiently familiar with these methods. They are being called upon to adapt a new assessment approach at the very same time that they are familiarising themselves with the principles behind this approach. An additional problem is that teachers do not perceive a clear connection between their teaching objectives and alternative assessment. They feel comfortable with traditional assessment. They believe that this type of assessment prepares their students best for the type of matriculation examinations in Israel.

This problem is interwoven with several considerations:

- Alternative assessment has just recently been identified as a basic principle of EFL teaching in Israel.
- EFL teachers have received little or no training in alternative assessment methods.
- EFL teachers lack the proper skills needed to create performance-based tasks.
- EFL teachers feel comfortable and secure with the traditional assessment methods they have used in the past.
- Alternative assessment methods require additional time considerations for which teachers are not being compensated.

- The planning of performance-based tasks calls for access to a wider range of resource materials, which may not be available.

### **1.3 STATEMENT OF THE PROBLEM**

In light of the above a need exists to investigate English foreign language teachers' understanding and implementation of alternative assessment in English foreign language. The key research problem can therefore be formulated as follows: How do EFL teachers incorporate standards and alternative assessment into their assessment practices? The following questions facilitate the demarcation of the problem:

- What are the theories underlying assessment?
- How do standards influence assessment?
- What does alternative assessment entail?
- What factors impact the teaching and assessment of English as a foreign language?
- Which assessment activities and methods do English foreign language teachers employ?

### **1.4 THE AIM OF THE RESEARCH**

This study focuses on assessment in teaching English as a foreign language in Israel. The study will explore teachers' understanding of alternative assessment and find out how teachers incorporate alternative assessment in their EFL classrooms. The aim is thus to:

- Find out how well EFL teachers in Israel understand the principles of standards and assessment
- Identify the types of alternative assessment currently being implemented
- Investigate teachers' attitudes towards alternative assessment

### **1.5 THE METHOD OF RESEARCH**

This thesis is an investigation into how the implementation of standards and alternative assessment methods has affected the teaching of English as a foreign

language in Israel. The research for this thesis will be comprised of both literary research and empirical study.

- The literary research will be a fusion of concepts of language learning, a comprehensive discourse of the theories framing foreign language study, a review of the history of the standards movement and its influence on various assessment methods. In addition, the principles and criteria for alternative assessment will be investigated, discussed and evaluated. The research will include an evaluation of traditional assessment methods, as well as a review and discussion of assessment alternatives.
- The empirical study will follow a qualitative design relying on textual analysis and focus group interviews.

According to Marshall and Rossman (1994:99) the choice of the data collection method depends on whether the method will provide adequate information, be cost effective and be feasible in terms of the sensitivities in the setting and the resources available for the study. It should also be linked in a logical way to the conceptual framework and research question and the overall strategy of the study.

The researcher chose to adapt a qualitative research design that was comprised of focus group interviews and textual analysis. Since the aim of the study was to investigate how EFL teachers understand and implement standards and assessment it seemed appropriate to adapt a research design that would investigate the participants' own perspectives on the subject (Popay, Rogers & Williams, 1998:346). Focus groups provide a wealth of information since they are homogeneous groups that, by nature, encourage discussion (Johnson & Christensen, 2004:185). Therefore, the major characteristic of focus groups is the insight and data that is produced by the interaction between participants (Morgan, 1997:12). This yields the type of information that is relevant and necessary in order to really understand the subject. Focus groups make it possible for the researcher to observe teachers in a natural setting where they can be comfortable to express their opinions and share their views

(Lofland & Lofland, 1995:18). The study will analyse the data collected from two focus group interviews.

Textual analysis is also valuable in an empirical study of this kind because it provides an understanding in context. The evidence in a textual analysis presents a clear and tangible record of the subject being investigated (Grady, 1998:24). The textual analysis in this study will collect, analyse and evaluate the assessment packages of ten English foreign language teachers. This type of investigation can provide significant data that can be beneficial in order to identify the assessment that is used by teachers in Israel. Both of these assessment instruments are valuable tools for investigating the depth and scope of a subject.

## **1.6 THE LIMITS OF THE RESEARCH**

The empirical study will be limited to a sampling of secondary high school teachers in Israel.

## **1.7 CLARIFICATION OF CONCEPTS**

Certain important concepts which recur in the text will be discussed in this section. This discussion serves as a mere introductory clarification of concepts. In the text more detail on the concepts will be provided.

### **1.7.1 Assessment**

Dietel, Herman and Knuth (1991:1) define assessment as “any method used to better understand the current knowledge that a student possesses”. Assessment refers to all those activities undertaken by teachers – and by the students in addressing themselves – that provide information to be used as feedback to modify teaching and learning activities (Black & William, 1998:140).

### **1.7.2 Traditional assessment**

Traditional assessment commonly refers to the tests that are administered to assess the acquisition of knowledge. These tests usually consist of select-response items, such as multiple-choice tests, true/false, fill-in-the-blank and matching exercises. Students are asked to recall or recognise information, rather than construct new knowledge. This type of assessment is usually carried out in a formal atmosphere, under controlled conditions. Critics fault traditional assessment methods with failing to tap the student's full range of abilities. Brown (1990:8) argues that traditional assessment neglects the assessment of practical skills, personal development, attitudes and performances in contexts outside the conventional classroom and laboratories. Her views are shared by Murphy and Torrance (1990:12) who complain that, "Assessment has been viewed for far too long as a formal process, which normally involves the administration of formal tests and examinations through procedures that are totally divorced from the educational process and setting to which they are supposed to relate".

### **1.7.3 Alternative assessment**

Alternative assessment (also known as performance assessment or authentic assessment) refers to strategies that are utilised to ask pupils to exhibit what they can do. "Alternative assessment refers to procedures and techniques which can be used within the context of instruction and can be easily incorporated into the daily activities of the school or classroom" (Hamayan, 1995:213). Huerta-Macias (1995:9) defines alternative assessment as situations in which "students are evaluated on what they integrate and produce, rather than on what they are able to recall and reproduce." He adds, "Although there is no single definition of alternative assessment, the main goal is to gather evidence about how students are approaching, processing, and completing real-life tasks in a particular domain". Pierce and O'Malley (1992:2) define alternative assessment as "a method of finding out what a student knows or can do that is intended to show growth and inform instruction and is not a standardised or traditional test".

#### **1.7.4 Assessment for learning**

Assessment for learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there (Assessment Reform Group 2002). This implies that assessment should not be seen as something that happens after learning; an add-on experience. It should rather be seen as something that promotes learning. It should be assessment *for* learning rather than just *of* learning (Gravett & Geysler, 2004:91).

#### **1.7.5 Feedback**

Feedback is information that supplies the performer with direct, usable insights into current performance, based on actual differences between current performance and desired performance (Wiggins, 1993:182). Sadler (1998:78) believes that feedback is concerned with praise for effort which will lead to higher self esteem, more effort and ultimately higher attainment.

#### **1.7.6 English as a foreign language**

English as a foreign language refers to the teaching of English where the setting of the class is usually monolingual and learners live in their own country (Kreiger, 2005:8). Mitchell and Myles (1998:1) and Brown (2001:16) define foreign language learning as the learning of a language that is not spoken in the local area and therefore, there is little or no exposure to the language.

#### **1.7.7 English as a second language**

English as a second language refers to the teaching of English where the setting of the class is usually multilingual and learners live in the culture of the target language (Krieger, 2005:5). There is high visibility and ready access to the target language outside the language classroom (Brown, 2001:116). Mitchell and Myles (1998:1) define second language learning as the learning of a language to any level under the

condition that the second language learning occurs sometime after the attainment of a primary language.

### **1.7.8 Standards**

Collins Cobuild English Language Dictionary (1992:1421.) defines a standard as “a level of quality or achievement, especially a level that is thought to be acceptable. It is something used to measure or estimate the quality or degree of something, for example, how good a piece of work is”. According to Marzano (Scherer, 2001:18), standards “identify the knowledge and skills that a student needs to know to be considered knowledgeable in a certain area and mandating the level of knowledge and skills that all students must achieve”.

In national discussions on the subject of education, standards have been defined as “what students need to know and be able to do, the essential core knowledge in a particular subject area, a passing score on an assessment, or a model demonstration representation worthy of emulation - much as an expert figure skater or diver ‘sets’ the standard” (O’Neil, 1993:4).

### **1.7.9 Language standards**

Language standards in regard to English refer to an agreement of a common linguistic plan that accepts the diversity of different variations of English despite the varieties of English that are spoken in all communities that use English (Brutt-Griffler, 1998:389). Widdowson (1994:385) refers to language standards as the stabilisation of the English of different communities into a standard form in order to retain the level of communicative effectiveness.

### **1.7.10 Standards movement**

The standards movement is primarily concerned with fairness and equality in education. Its aim is to replace the "tracking and sorting" objectives of school systems in the past with equitable educational opportunities. The movement is

directed towards maintaining high expectations for learning and providing support for all students, teachers and educational leaders (Thompson, 2001:358).

### **1.7.11 Task**

A task is a “goal-directed activity, demanding that the student use his background of knowledge and skill in a continuous way to solve a complex problem or question” (Dietel, Herman & Knuth (1991:1). The latest trends in assessment advocate the incorporation of tasks that are performance-based; activities that involve some kind of presentation before an audience.

### **1.7.12 Performance-based tasks**

Performance-based tasks are activities such as projects and exhibitions, which call upon students to apply their knowledge, skills and understanding in important, real-world contexts.

## **1.8 PLAN OF STUDY**

The following program will be followed:

- **CHAPTER 2**

Chapter 2 will be a theoretical chapter that will concentrate on the subject of assessment. The literary study will investigate the development of assessment, and will explore and review various assessment theories. In addition, it will trace the transition from traditional assessment practices to alternative assessment applications, giving special attention to alternative assessment methods that are most suitable for assessment in foreign languages. The chapter will include discourse as well as debate of theories and applications.

- **CHAPTER 3**

Chapter 3 will deal with the implementation of standards in education and its impact on assessment practices. Special attention will be given to a discussion of the effect of the standards movement on teaching and assessment in Israel.

- **CHAPTER 4**

A theoretical study will comprise the bulk of chapter 4 which will deal with the instruction of English in the education system. The chapter will explore the theories and practices that have dominated language education and the principles that have formulated these theories. In addition, the chapter will discuss the changes and transformations in the approach to language instruction in Israel, which led to new forms of assessment and the implementation of performance-based instruction.

- **CHAPTER 5**

Chapter 5 will contain the research design and explain the significance of a qualitative research plan.

- **CHAPTER 6**

Chapter 6 will include the results of the investigation carried out in the previous chapter. Conclusions will be provided based on the findings of the investigation.

- **CHAPTER 7**

Chapter 7 will discuss the educational implications of the empirical study. The chapter will present recommendations.

## CHAPTER 2

### FROM TRADITIONAL ASSESSMENT TO ALTERNATIVE ASSESSMENT

#### 2.1 INTRODUCTION

Regular assessment occurs daily in the classroom, where teachers listen to and observe their students on a steady basis. They are able to sense with relative accuracy, which students seem to understand and which students appear confused. Assessment is a valuable tool for teachers and students for a myriad of reasons. Firstly, teachers need to assess their students in order to monitor their progress. Teachers want to know that their students are paying attention and that they are on task. In addition, it is important for teachers to ascertain that their students have mastered material that has previously been taught before they advance to more comprehensive and complex subjects. Secondly, assessment is an effective diagnostic instrument for identifying individual students who need additional help, but are either too uncomfortable or embarrassed to seek assistance. This implies that not only the lack of mastery of objectives are identified, but difficulties in the process of mastery are also identified in the process. Thirdly, in addition to determining whether the learner passes or fails, assessment also has an influence on possible admission to university study or certification.

Students themselves want assessment for several reasons. Primarily, they look for a tangible means of gauging their own achievement, and usually regard the grade on a paper or examination as a reflection of their achievement.

Assessment practices have perpetuated the understanding that a passing grade affirms success, while a failing grade shows the inability to measure up to desired standards. In addition, assessment offers students a form of closure at varied stages of learning. They expect a test at the end of a unit so they can feel a sense of accomplishment before they continue on to other more intricate tasks and objectives. Assessment enables them to feel proud that they have mastered a subject or sounds the alarm that they are not quite up to par.

Formal assessment, primarily in the guise of testing has become a bastion for educators who relish the scientific results that such evaluation yields (Allan, 2000:26). Written tests that are numerically scored have become the standard that teachers rely on. They, their students and parents easily relate to numerical grades and their universal meaning. Formal assessment in the form of tests has been widely used for accountability; to measure and compare student achievement on municipal, state, national and global levels. Educational ministries and institutions depend on the results of assessment in order to evaluate educational policies. They make curriculum decisions, create policies, distribute material, allot funding, and hire staff based on the results of test scoring data (Hancock, 1994). Policy makers frequently offer incentives, such as increased funding and additional benefits, for schools that produce high scores on examinations and sanctions on schools that do not (Stiggins, 2002:759; Popham, 2003:123).

Testing is also used to separate the better achievers from less successful achievers. As a result, the educational system has come to depend on wide-scale testing for admittance and rejection purposes. However, the gross misuse of assessment for selection has recently come under attack. The reason for the dissatisfaction is that traditional tests are in reality a poor judge of what students really know (McTighe, 1997:6) and they fail to measure many of the important facets of learning (Darling-Hammond, Ancess & Falk, 1995:4). An additional problem is that teachers teach to the test as opposed to the skills or the wider curriculum. The greater issue is whether large-scale assessment should occupy such a pivotal role in accountability or in facilitating or retarding student advancement (Stiggins, 2001:5). Shouldn't most time, effort and resources be dedicated to assessment in the classroom, where actual learning takes place regularly? These questions and issues warrant a thorough investigation into what assessment entails and the principles of assessment. This chapter explores the theories underlying assessment procedures and endeavors to provide an acceptable definition of assessment and identify the principles that influence assessment.

## **2.2 THEORIES UNDERLYING ASSESSMENT PRACTICES**

Assessment methods changed over time to coincide with the findings of more current

studies in the area of knowledge acquisition. As researchers uncovered new information on how the mind works educators were called upon to adapt teaching methods to meet the new needs of students. Teachers were directed to enlist activities that would appeal to students' interests, learning abilities and learning styles. Consequently, assessment procedures have had to change in order to best evaluate the different approaches to learning, and to incorporate principles based on the newest theories. The most prominent theories that influenced assessment and led to changes in assessment were: behaviorism, cognitivism, constructivism and multiplicity of intelligence.

### **2.2.1 Behaviorism**

Behaviorists dominated the first half of the 20<sup>th</sup> century with their belief that a person's actions were completely based on stimuli or random occurrences in the environment. Behaviorists maintained that learning could be maximised by the accumulation of small bits of knowledge, and by rigidly organising information in sequential and hierarchical order (Mitchell & Myles, 1998:23; Shepard, 2000:5). This idea of dividing processes into numerous, small steps and reinforcing each step frequently and systematically in order to achieve maximum learning was pioneered by B.F. Skinner in the 1930s. Since he believed that people "are nothing more than simple mediators between behavior and the environment", he directed the focus of learning on the external environment and the ways in which it shapes behavior (Skinner, 1931:428).

As a result of Skinner's research, behaviorists promoted a mechanical approach to learning, which consigned the learner to a relatively inactive role. Therefore, teaching was based on factual and procedural rules that students were required to follow. Study material was broken down into small instructional steps, and memorisation, drilling and rote learning were considered effective methods of learning. The behaviorist views heavily influenced assessment.

As motivation was externally generated and learning was based on the reinforcement of numerous small steps, the "one-skill-at-a-time" (matching, multiple-choice, true/false) test items were popular with behaviorists (Shepard, 2000:5). Tests that

offered questions like multiple-choice, true/false, matching and fill-ins could check the distinct skills that students needed to master by relying on responses to external stimuli (i.e. "choose the correct answer from the choices below") rather than requiring students to develop their own streams of thought, a skill which they were not trained to develop. In addition, this type of testing was appealing since it is relatively quick and efficient to administer (Worthen, 1993:446).

The subsequent interest in constructive and cognitive approaches to learning changed the direction of language instruction; turning the view of learning away from the environment and inward toward the student.

### **2.2.2 Cognitivism**

Whereas the 1940s and 1950s abounded with behaviorist views on learning, the following years marked a rejection of the previously accepted behaviorist theories, which were now believed to limit the individual's potential for knowledge acquisition (Alvermann & Phelps, 1994:36). The cognitive revolution, which gained recognition in the 1960s, emphasised the importance of the mind in the learning process. Although repetition and reinforcement were still considered important, attention was given to how the individual processes and stores information. Learning is seen as "an active process of mental construction and sense making" (Shepard, 2000:6). It is important to mention here that cognitivism and cognitism are the same.

Cognitive psychologists maintain that cognitive learning principles influence teaching and learning. The main objective of learning is being able to apply skills and knowledge in order to perform meaningful tasks in authentic situations (Wiggins & McTighe, 2006:26). In addition, cognitive abilities are created and strengthened through socially supported interactions. Therefore, it is understandable that as language began to center on its communicative role, more emphasis was placed on the importance of social interaction and performance in authentic situations connected to the real world (Shepard, 2000:7). It was believed that students who desire to communicate successfully would invest time and exert effort to perform well.

Cognitivism contends that teaching and learning are influenced by cognition, which Atkins (1993:257) describes as the “symbol-processing view of the mind”. He explains that true learning takes place when there is a direct connection between stimulus and response. Since each person relies on his own abilities and own experiences to synthesise the information that he is exposed to, we can understand how two similar students react differently to the same lesson, why one student grasps the material being taught and why another seems to exhibit little or no comprehension (Gagne, Yekovich & Yekovich, 1993:4-5). This information is important in helping teachers decide what methods should be used to teach a given topic. In addition, this insight suggests that an individual approach to teaching, learning and assessment can yield more productive and effective results.

Cognitive psychology also provides some insight regarding a student’s desire to learn as well as the mental processes that support problem solving. Cognitive psychologists claim that learners have “cognitive data banks” pre-stocked with empirical knowledge or methodological rules, which fuse an understanding between the outside world and human thought (Philips, 1995:5). Therefore, cognitive psychologists insist that the behaviorist approach to learning is incorrect. Students need to process information and relate it to their own experiences. Assessment needs to extend beyond factual recall and selecting a correct response from a choice of possible answers. Assessment should reflect understanding, reasoning and problem solving (Shavelson & Baxter, 1992:38).

Cognitivism has had a considerable impact on assessment. Since cognitive learning suggests learning on a more complex scale than memorisation and recall, teaching and assessment in the light of this understanding is not served by traditional testing methods. Cognitive theories have brought attention to the concept of the mind, and the understanding that learning is “an active process of mental construction and sense making” (Shepard, 2000:6). Learning, therefore, needs to be assessed by methods that call for greater student activity and involvement. Alternative assessment requires students to integrate and produce (Huerta-Macias, 1995:9) and is more suited to cognitive learning.

Alternative assessment can be the vehicle through which students show what they can do with their knowledge and skills. They can provide scenarios where students can transfer learning between contexts and carry information from one area of study to another.

### **2.2.3 Constructivism**

While cognitivism advanced the area of educational psychology, constructivism developed these theories even further in the 1990s. Constructivism, like cognitive psychology firmly places the individual at the center of the learning process. The theory of constructivism argues that human beings do not learn by passively absorbing knowledge. Instead, they are active participants in the acquisition of knowledge, building their own understandings, drawing on prior knowledge and formulating mental concepts (Perkins, 1992:77). This explains why different learners construct different meanings from learning experiences since they call on their own individual backgrounds. Teaching principles during this period supported activities that engage students in issues and experiences that are familiar, where the focus of learning is on interpreting meanings, discovering and problem solving (Harris, 2000).

A feature of constructivist teaching is that it is more student-centered, and it requires students to become active participants in the learning and assessment processes. Students who are involved in their own learning experience by making their own associations and creating their own understandings tend to be more motivated and anxious to succeed. Teachers, however, have to adjust their teaching practices to allow their students to assume a more active role. They need to cultivate a safe, encouraging environment where students will feel comfortable to express their own ideas and develop their own concepts (Airasian & Walsh, 1997:448).

Alternative assessment methods in the form of presentations, projects, portfolios and creative expressive activities (such as journals, diaries and student logs) offer students opportunities to demonstrate what they are able to do with the knowledge and skills they have acquired. It challenges them to create and express themselves by drawing upon the host of experiences they have shared.

### 2.2.4 Multiple intelligences

Another principle that has revolutionised the area of educational psychology is the theory of Multiple Intelligences (MI) developed by Howard Gardner in 1983. Gardner's theory of MI provides a theoretical foundation for recognising the different abilities and talents of students. Approaching and assessing learning in this manner allows a wider range of students to successfully participate in classroom learning (Brualdi, 1996).

This theory proposes that there are seven (he added an eighth intelligence - naturalist intelligence - in 1995) different aptitudes worthy of being designated intelligences. Gardner proposes that an individual has a number of specific types of intelligence in varying degrees, and that instruction should be related to the student's individual needs (Gardner, 1993:17). These distinct human intelligences, skills or aptitudes include: 1) musical, 2) bodily kinesthetic, 3) spatial, 4) interpersonal, 5) intra-personal, 6) natural, 7) logical mathematic and 8) linguistic. According to Gardner, only the latter two intelligences: logical mathematics and linguistics are regularly tested and appreciated in schools. However, he believes that the eight intelligences are socially and culturally relevant in society and, therefore, should be addressed in the classroom.

Gardner (1995:208) justifies his belief in MI on the principles that not all people are the same; people think differently and, therefore, a single educational approach cannot possibly serve everyone. Gardner's views are valid and schools should develop programs that will identify the areas in which children show the greatest strengths and direct teachers towards those strengths (Hatch, 1997:26). He suggests that rather than question the degree of intelligence that each young child possesses (i.e. intelligence testing) a better consideration would be to examine the ways in which the child exhibits intelligence and to offer opportunities where he can successfully demonstrate his abilities.

Gardner (1994) not only invalidated the belief in a single kind of intelligence, but also the previously held view that intelligence can be sufficiently measured in a short period of time. According to Gardner, learning and teaching should focus on the

particular intellectual abilities of each individual, which, most importantly, should be monitored over time.

Multiple intelligences, constructivism and cognitivism share the belief that individuals relate to stimuli and challenges by drawing upon a stable reservoir of abilities and experiences. As educators began to embrace these new theories, the trends in education shifted to meet the new demands of alternative approaches to learning and assessment. Although traditional assessment had retained prominence throughout the period when behaviorist theories were popular, in the light of more current research this form of evaluation was no longer considered to be relevant or productive. As a result of this understanding, new thought was given to discussing and deciding what were the purposes and principles of assessment and how teachers could most effectively evaluate their students.

MI has had a significant influence on assessment practices due to the understanding that learners need to be offered a variety of opportunities to demonstrate their knowledge and skills and to allow for varied approaches to learning (Ghosn, 1997).

Armstrong (1994:115) maintains that:

“MI theory proposes a fundamental restructuring of the way in which educators assess their students’ learning progress. It suggests a system that relies far less on formal standardised or norm-referenced tests and much more on authentic measures that are criterion-referenced, benchmarked or ipsative (i.e. that compare a student to his or her own past performances).”

Now that it is clear which theories underlie current assessment practices, it is necessary to look in more detail at assessment, as such.

### **2.3 DEFINING ASSESSMENT**

In order to conduct a study on assessment, it is necessary to have a clear understanding of the meaning of the word. When we talk about assessment in education today we refer to the evaluation and appraisal of student performance. In traditional assessment methods, which were widely used in the past, teachers assessed students essentially at the culmination of learning activities. Moreover, the

responsibility of assessment was delegated to the teacher leaving the student relegated to a passive role.

Champions of alternative means of assessment, who have a very divergent view of assessment, believe that true assessment is contingent on a reciprocal relationship between assessor and assessee (Wiggins, 1993:14). They define assessment as the mutual interaction between assessor and student who work together in order for students to reach maximum learning capacity.

Wiggins (1993:14) traces the etymological origin of the word “assess” to a form of the Latin verb “assidere” which means “to sit with”. He explains that in assessment, one “sits with” the learner, and “assigns” value similar to the service provided at one time by the tax assessor. The assessor attempts to bring out all that a student knows and can do by various means. His responsibility is not merely to record student responses, but rather to act as a facilitator enabling the student to draw on his skills and knowledge in order to produce results (Wiggins, 1993:14).

According to Cotton (1995:89), assessment is an instrument used to describe a person’s level of achievement or potential. Hancock (1994) shares related thoughts about assessment, which he views as an ongoing strategy through which students’ learning is monitored. He, however, emphasises the importance of students being involved in making decisions about the extent to which their performance matches their ability. Rust (2002) similarly defines assessment as an evaluation or appraisal, which is about making a judgment. Assessment, he believes, identifies strengths and weaknesses, good and bad, and right and wrong in some cases. He believes that assessment is more than simply giving marks or grades, although that may well be a part of it. Assessment “involves collecting evidence of learning over a period of time, using a variety of assessment methods” (Ministry of Education, Culture and Sport, 2001:19).

The definitions mentioned above share the common understanding that the purpose of assessment is to consider a student’s educational attainments. In most cases assessment involves collecting and interpreting, which result in decisions about a student’s abilities.

### **2.3.1 Assessment and evaluation**

Assessment and evaluation are often used interchangeably although they are different. Assessment is a general term for all efforts to obtain information pertaining to learners' performance and ability, irrespective of purpose (Todd, 2002:16). Airasian (1997:3) cites the various elements of assessment by defining it as "the process of collecting, synthesising and interpreting information to aid in decision making". Sieborger and Macintosh (2004:5) believe that assessment describes a broad array of methods that are used to measure a learner's accomplishment.

Evaluation, on the other hand, refers to something bigger than assessment. Sieborger and Macintosh (2004:5) view assessment as just one aspect of evaluation. Evaluation entails judging or deciding about the value and worth of a performance or activity (Airasian, 1997:4). It is used for those efforts that generate quantitative data that can be used to produce scores measuring the learner's performance and ability (Todd, 2002:16).

Clearly, both assessment and evaluation must be considered when viewing student performance. Whereas the former stresses the importance of measuring progress over time, the latter recognises the value of the student's ultimate accomplishment.

There are several forms of assessment, which are differentiated by their role and purpose in the learning process.

### **2.3.2 Testing and assessment**

There seems to be a misunderstanding of the terms testing and assessment. Since these two words are so frequently interchanged with each other they are often thought to have identical meaning. Yet there are distinct differences between the two words, and their intentions are different. There are a variety of means by which students' abilities are measured. The teacher normally makes a decision on the grounds of information obtained from measuring by means of tests and examinations, which are popular measurement instruments. Tests are a "formal, systematic, usually paper and

pencil procedure to gather information about pupils' learning" (Airasian, 1997:3). Tests are only one of the many types of assessment information that teachers utilise.

Allan (2000:27) aptly calls tests "snapshots", which are brief moments in the continuous learning process, while he regards assessment as "a set of processes that go on the whole time". A single grade or score generally reports the results of tests, and the test grade can influence a decision to retake the test, pass or fail a course or proceed to the subsequent unit. However, whereas testing focuses on measuring accomplishment at particular intervals, assessment should include a wide spectrum of information that teachers collect in their classrooms that assist them in understanding and supervising their pupils' accomplishments (Airasian, 1997:5).

Although the monitoring of learning is a common feature of both testing and assessment, it is apparent that the latter extends far beyond the scope of measuring achievement. As the objectives of assessment are to improve, develop, and draw a parallel between ability and performance, assessment is a more pervasive activity (Eisner, 1999:658). The assessor needs to show patience, consideration and a willingness to provide continuous guidance (Wiggins, 1993:14). The test giver, on the other hand, is often detached from the students. This might be explained by the fact that the testing process, by its very nature, segregates teacher and student. In a testing situation the students are seated apart from the tester in an environment where communication between them is either discouraged or forbidden. In some test situations, such as standardised testing and achievement testing, the students have never met the teachers who administer the test, which results in feelings of uneasiness and apprehension.

From the above it is clear that assessment is more expansive than testing and as a result the role of the assessor becomes more comprehensive than the role of the tester. This is apparent in the work they do. Whereas the tester checks answers to questions that he has previously chosen, the assessor must be prepared to respond to and direct areas of performance that he might not have anticipated. While the tester spends a limited, specified time on evaluation (constructing the test and marking it), the assessor's work is spread over a longer period since assessment monitors, documents, and assesses work in progress. Considering the nature of their responsibilities, it is

clear that the assessor faces greater challenges and therefore, must be prepared to meet them.

Perhaps the most significant difference between testing and assessment is their intended objective. High stakes testing, such as standardised testing and achievement testing is concerned with comparing student achievement generally for ranking purposes. Assessment, on the other hand, is used to strengthen individual student progress (Stiggins, 2002:761).

The difference in perceptions of testing and assessment are just one of the factors that have influenced changes in assessment procedures. The major influence has been the evolution of theories relating to how the mind acquires and processes knowledge. The principles that outline these theories have brought about the transition to new, preferred methods of alternative assessment.

### **2.3.2.1 Standardised testing**

Standardised testing refers to any assessment device that is administered and scored in a standard, predetermined manner (Popham, 2003:125). Standardised testing first became popular in the 1960s as a method of checking school progress on a district wide level and was expanded in the 1970s to include statewide evaluation. The belief that standardised testing was the panacea that would raise the academic level of all students led to considerable financial investments in nationwide assessment in the 1980s and quickly extended to international standardised testing in the 1990s (Stiggins, 2002: 759). An important feature of standardised tests is the emphasis on speed. Students are expected to complete numerous questions in a strictly limited time frame. The “speed” format allows for a widened distribution of scores (Rogers, 1995:256).

### **2.3.2.2 Achievement testing**

Achievement tests, which began to be used in the early 1920s, were initially intended to measure the knowledge and proficiency that students learned in school (Popham, 2001:46). However, they quickly became a vehicle for comparing test takers. These

types of tests are unfair because many questions are based on knowledge acquired outside the classroom. Therefore, they are biased against students of lower socioeconomic groups (Popham, 2001:46). In addition, this researcher believes that performance on these tests is heavily influenced by language skills, which discriminates against particular student populations.

An additional problem of both standardised and achievement testing is that they require students to respond independently to pre-ordained questions, which cannot be clarified. Furthermore, students receive final grades on tests without having the opportunity to explain or extrapolate their answers. As a result, standardised and achievement testing instead of raising the potential for learning, have caused many students to feel discouraged and inadequate (Stiggins, 2002:758).

### **2.3.2.3 Aptitude testing**

Aptitude testing is intended to predict a student's future performance, typically in a succeeding academic setting (Popham, 2003:45). This type of testing is designed to show students' general ability to learn and/or to achieve competence (Brown, 2004:43).

### **2.3.3 Types of assessment**

There are several types of assessment that are utilised at different times and for different purposes.

#### **2.3.3.1 Baseline assessment**

Baseline assessment is usually done at the start of a grade, but can also occur at the beginning of any learning cycle. It is used if teachers would like to determine what the learners already know and can do before they commence teaching at the beginning of a year. Baseline assessment enables teachers to determine learning deficits early on and adapt teaching strategies accordingly (Fraser, Loubser & Van Rooy, 1993:206).

### **2.3.3.2 Diagnostic assessment**

Diagnostic testing is designed to diagnose specific features in an area of study. For example, in diagnostic language testing, the test might diagnose the phonological characteristics of English that are difficult for the student (Brown, 2004: 46). These tests assess learners' strengths and weaknesses and often are used for remedial purposes (Kilfoil & Van der Walt, 1997:290).

### **2.3.3.3 Formative assessment**

Formative assessment is assessment that is intended to ascertain what progress a student is making during learning and to provide feedback on it (Cotton, 1995:24). It means assessment “that is specifically intended to provide feedback on performance to improve and accelerate learning” (Sadler, 1998:77).

Airasian (1997:126) points out the distinguishing features of formative assessment, such as:

- It monitors and guides a process while it is still in progress.
- Assessment takes place during the process.
- The use of assessment information is to improve and change a process while it is still going on.

Clearly formative assessment is in tune with the understanding that assessment monitors work in progress with the intention of improving student achievement. Assessment of this kind can complement learning and guide students towards greater accomplishment.

Continuous assessment, which is assessment that occurs “on and off” throughout a time of learning, should be used for formative purposes (Sieborger & Macintosh, 2004:23).

Continuous assessment is beneficial for several reasons (Sieborger & Macintosh, 2004:23):

- It requires teachers to think about assessment throughout the learning process.
- It facilitates the implementation of an outcomes-based curriculum.
- It provides learners with greater opportunities to take control of their own learning
- It enables teachers to assess a diversity of learners according to differences in rates and styles of learning.

Stiggins (2002:761) agrees that assessment for learning purposes can best be achieved through continuous assessment.

#### **2.3.3.4 Summative assessment**

The distinguishing feature of summative assessment is that it occurs at the end of a program (Cotton, 1995:24). Other characteristics of summative assessment, outlined by Airasian (1997:126) are:

- It judges the success of a process at its completion.
- The time of assessment is at the end of a process.
- The use of assessment is to judge the overall success of a process and to give a grade.

Clearly, summative assessment is concerned with the definitive and definite outcome of student performance. As attention is not given to works in progress nor in providing constant and continuous guidance, it is obvious that this type of assessment contributes little to ensuring successful learning outcomes.

#### **2.3.3.5 Ipsative assessment**

Ipsative assessment measures individual improvement by comparing the grade or level at the start and the finish of a learning programme (Cotton, 1995:27). This type of assessment, which also is the basis of qualitative research, has been ignored in favor of other types of assessment. However, greater academic advancement can be attained not by forcing students into competition with each other, but rather by stressing the importance of the individual's own academic development and progress.

### **2.3.3.6 Norm-referenced assessment**

Norm-referenced assessment is a test or other kind of assessment created to give an appraisal of a performance that is interpretable in terms of an individual's comparative rank in some known group (Linn & Gronlund, 2000:42). Often in this type of assessment, easier questions might be eliminated and more difficult questions added in order to facilitate ranking (Linn & Gronlund, 2000:43). Furthermore, norm-referenced tests often have strict, unreasonable time limits in order to further prevent a wide range of students from being successful.

A problem with this type of assessment is that there are a number of reasons, such as good or poor teaching, that might affect an individual's performance, which would result in an inaccurate evaluation (Rust, 2002:2). In addition, a common feature of norm-referenced assessment is the curve of "normal distribution" which presumes that a certain number of applicants will do extremely well and a few will do poorly and the majority will perform in the middle as average (Rust, 2002:2). Therefore, it is possible that test results might be manipulated in order to fit the intended outcome (Rust, 2002:2).

### **2.3.3.7 Criterion-referenced assessment**

In contrast, criterion-referenced assessment is a test or type of assessment created to give an evaluation of a performance that is interpretable in terms of a clearly defined and "delimited" area of learning tasks (Linn & Gronlund, 2000:42). Criterion-referenced assessment differs from norm-referenced assessment in that it is concerned with learning outcomes that can provide information that will improve instruction (Linn & Gronlund, 2000:43). Particular abilities, skills or behaviors are each specified as a criterion, which must be reached (Rust, 2000:2).

Linn and Gronlund (2000:43) suggest three ways in which teachers can make criterion-referenced analyses. For example:

1. Describe the particular learning tasks a student is able to perform (e.g. counts from 1 to 10).

2. Show the percentage of tasks a student performs correctly (e.g. spells 65 percent of the words in the word list).
3. Compare the test performance to a set performance standard and decide whether the student meets a specified standard (e.g. performed at the proficient level).

Clearly, criterion-referenced assessment is in tune with the growing understanding that the potential value of assessment goes beyond testing, particularly high stakes and accountability testing. Unlike norm-referenced testing, which does little to enhance the learning potential of individual students, criterion-based assessment contributes to the student's productivity.

### **2.3.4 Measurement and measurement instruments**

Any assessment method depends on measurement and reliable measurement instruments.

Measurement is the process of quantifying or allocating a number to a performance (Airasian, 1997:3). It is the methodical distribution of numbers or names to features that are under study (Hanna & Dettmer, 2004:5). A measurement instrument measures by utilising a sophisticated device or systematic procedures (Hanna & Dettmer, 2004:5). A test, for example, is a formal systematic procedure to gather information under uniformed conditions (Hanna & Dettmer, 2004:5). Traditional assessment tests rely on scoring questions according to point values that are decided upon prior to testing.

Alternative assessment tasks require different measurement instruments that are suitable for judging performance tasks and projects. Mere numerical grading cannot efficiently or effectively assess more complex and multifarious activities. Evaluation based on clear and explicit performance criteria are needed. The most commonly used measurement instruments for alternative assessment are scoring guides, such as checklists and rubrics. They describe what to look for in products or performances to judge their quality (Arter, 2000).

#### **2.3.4.1 Checklists**

Checklists are a written list of performance criteria that students are supposed to reach (Airasian, 1997:226). They are a valuable addition to any assignment since they provide clear and explicit direction for students. There are innumerable items, which can be included on a checklist, such as: a) I included a cover page and table of contents, b) I wrote the bibliography correctly, c) I asked five grammatically correct questions to ten different people for my survey, and d) I checked for language and spelling mistakes. The decision of which items to include should depend on the objective of the assignment as well as on the skills that are being checked. Checklists must be given to the students prior to commencing an assignment in order to be effective so that they know what they are expected to do. Once a student's performance is completed, a checkmark is recorded next to those criteria that have been successfully met. Checklists can empower students with self-confidence and motivation by affording them with specific and clear criteria.

#### **2.3.4.2 Rubrics**

Rubrics is an assessment tool that verbally describes and scales levels of student achievement or performance tasks, but it can also be associated with more conventional alphanumeric and numeric scores and grades (Solomon, 2002:58). Rubrics describe the performance behaviors that mark the intervals or comparative characteristics and provide evidence of achievement (Solomon, 2002:58).

There are fundamentally two uses for rubrics (Arter, 2000). Firstly, they collect information on students in order to plan instruction, trace student progress toward important objectives, and report progress to others. Secondly, rubrics help students become progressively more competent on the performances and products being assessed by providing guidance and direction (Arter, 2000).

Popham (2003:96) outlines several important components in any rubric:

- Evaluation criteria - refers to the factors that a scorer considers when establishing the quality of a student's response. For example, if the

organisation of an essay is a factor there must be an essay-scoring rubric.

- Quality definitions - explain what is needed in respect to each evaluation criteria for a student to receive high versus low rating. For example, “essay organisation” would identify types of organisational structures that are acceptable and those that aren’t.
- Scoring strategy - indicates how a scorer should use evaluation criteria and their quality definitions in two main ways: holistic and analytic.
  - a) holistic - implies that the scorer must address the degree to which a student’s response complies with all the evaluation criteria in order to provide a broad, general evaluation.
  - b) analytic - constructs a criterion-by-criterion judgment for each evaluation criterion via predetermined possible numbers that can be combined to reach a final score.

Numerical summarisation is the most commonly used and most basic rubric (Airasian, 1997:228). This method allocates point value to each category in a given scale and sums the points across performance criteria. Another form is descriptive summarisation based on short, written descriptions of different levels of performance, for example: excellent/good/ fair/poor (Airasian, 1997:231).

In order for rubrics to be effectual, there must be a limited number of rating categories to choose from. Airasian (1997:228) suggests that threes to five well-defined rating points are optimal.

Rubrics and checklists are different than traditional assessment instruments in that they use criteria to increase the quality of student performance, rather than simply evaluate it. In light of findings, which cite the role of assessment in learning, rubrics and checklists, are effective assessment and learning instruments.

## **2.4 PRINCIPLES OF ASSESSMENT**

The search for more effective ways of monitoring student progress (section 2.2, *Theories Underlying Assessment Practices*) became closely connected to a

reconsideration of the purposes and principles of assessment. As the objective became to help students learn, rather than to check how much they have learnt, several factors were considered to be very important. First, it was understood that instructions had to be clear and explicit so that students could understand what was expected of them. Second, for assessment to be really accurate and reliable it needed to be measured over a period of time. Therefore, formative assessment (assessment provided during learning) rather than summative assessment (assessment given at the end of the learning process) gained prominence. In addition, it was also understood that for formative assessment to be most beneficial it must provide feedback that students can use to improve their performance (Rust, 2002:1).

Airasian (1997:149) adds three additional features that could certainly be applied to both traditional and alternative assessment measures.

1. The assessment exercises and questions are correlated to the teacher's goals and instruction.
2. The exercises and questions cover material that is comparable to what was taught in class.
3. The items, instructions, and scoring processes are explicit and suitable.

In addition to the factors mentioned above, there are a number of requirements that measuring instruments must comply with in order to ensure accurate assessment.

#### **2.4.1 Reliability**

In order for any assessment to be successful it must be reliable. Reliability means that any judgments or conclusions drawn from assessment should give a dependable picture of a student's work (Sieborger & Macintosh, 2004:12). A reliable test should produce similar results if the instrument is used a second or subsequent time either for the same or different learners (Kilfoil & Van der Walt, 1997:285; Cotton, 1995:93).

#### **2.4.2 Validity**

In order for assessments to be effectual they must be valid. Validity is defined as the degree to which a test actually measures what it is supposed to measure (Kilfoil &

Van Der Walt, 1997:284; Sieborger & Macintosh, 2004:11). A method of assessment is considered to be valid if “it measures the intended aims, goals, objectives, performance or quality” (Cotton, 1995:93).

For evaluation to be valid the evaluation should match the course objectives and parallel the material taught in class (Kilfoil & Van Der Walt, 1997:284). Todd (2002:17) is in agreement stating that, if for example, the objectives include increased motivation, positive attitudes towards English, and greater independence and awareness then self-assessment should be a serious consideration.

The validity of a test may be ascertained in a variety of ways.

- Content validity means carefully checking the test’s content in relationship to the syllabus.
- Concurrent validity means deciding the degree to which test results correlate with results of existing tests that have been validated.
- Predictive validity means checking the extent to which assessment results can be used to predict student performance (Kilfoil & Van der Walt, 1997:285).

For scoring rubrics to be valid, criteria should make it possible to effectively and convincingly differentiate performances of varied levels of quality (Wiggins, 1993:238). The discrimination between various degrees of performance should not be arbitrary. Wiggins (1993:239) argues that designating 6 as “good” and 5 as “average”, for example, is not a valid assessment. Rather he suggests that before scoring rubrics are created it is important to get a sample of different excellent performances and an example of a scope of possible performances so that actual student work can be assessed in relationship to these ranges. In order to grant validity, a measuring instrument should provide for objective appraisal (Lien, 1980:35).

### **2.4.3 Efficiency**

In addition to being reliable and valid, a test must be efficient. This means that it should be economical, easy to administer, score and interpret (Bachman, 1990:34).

#### **2.4.4 Transparency**

Effective assessment, in addition, must be transparent. Transparency in assessment means the degree to which all parts of the assessment process are clear and available to the student. Wiggins (1993:229) insists that “demystified criteria and standards” are an integral element of any assessment procedure. Assessment can be transparent if tasks are “discussed, clarified, and even appropriately modified through discussion with the assessor, and or one’s colleagues”. In order to achieve validity and reliability, the terms of learning outcomes and criteria must be clear and exact. Most important these terms have to be carefully outlined when the task is assigned so as to provide the maximised conditions for successful performance (Todd, 2002:2).

Transparency is very vital since understanding the foundation on which performance will be judged facilitates the improvement of performance (Linn et al., 1991:17).

Attention should also be given to assure that the assessed task is a valid measurement of the learning objective. These considerations are additional reasons for the shift from traditional to alternative means of assessment, which are believed to adhere more closely to the purposes and principles of effective evaluation. The following sections will discuss traditional assessment, which was popular for a significant period of time, and explain the factors that led to the implementation of alternative means of assessment.

### **2.5 TRADITIONAL ASSESSMENT**

Traditional assessment refers to a variety of testing procedures, such as: multiple-choice, true/false, matching and fill-in questions. These traditional assessment tests dominated the assessment scene for the last five decades probably because of the ease in which they could be administered in a wide variety of situations (Worthen, 1993:452). These tests allow a large number of people to be checked on identical material, at the same time, under similar conditions. They are relatively simple to create, easy to administer, reusable and uncomplicated to mark as they are usually answer-keyed or machine-scored. This litany of reasons explains why this type of

testing became popular and was adopted for standardised and accountability testing purposes.

Great weight was placed on traditional tests in the past as they were thought to be more reliable and valid than teacher assessment (Allan, 2000:26). Teacher assessment refers to all the classroom procedures that teachers engage in on a daily basis to check their students' progress, such as: asking questions, checking homework, marking classroom assignments and reports, and carrying out discussions.

Teachers' assessments of learners' performances was thought to be necessarily subjective and lacked the reliability of scientific evaluation. Whereas exams could be standardised, teachers' assessments could be prone to bias, favoritism and other factors (Wiggins, 1993:3). Moreover, educators felt more comfortable with clear, straightforward responses, which could be answered and marked objectively. They felt comfortable relying on tests, which loyally followed the "worldwide historical tradition of using figures to express test results" (Allan, 2000:26). These numerical scores, when necessary, could be easily transcribed with verbal meaning that could be clearly understood.

### **2.5.1 Types of traditional assessment**

Although traditional tests can feature a variety of questions, these questions have similar characteristics. The limitations of these types of questions are often overlooked in favor of their myriad of advantages.

- **Multiple-choice questions**

The multiple-choice question, which is the most popular test format for measuring achievement, involves matching a question or partially complete sentence with the correct answer from three or four possible options (Popham, 2003:81). Multiple-choice questions, however, can be a less accurate measurement of achievement since studies have shown that there is some element of random selection, which can result in successful answers (Airasian, 1997:155). This type of question also allows

students to recognise correct answers that they might be unable to produce on their own; thereby presenting an exaggerated representation of a student's understanding (Popham, 2003:81). An additional disadvantage of multiple-choice questions is that they fail to develop creative thinking (Popham, 2003:82) or require students to use any productive language (Brown and Hudson, 1999:685).

- **Short-answer questions**

The short-answer question is the easiest to construct and measures recall of information. The limitations of this type of questioning are that they are not suitable for measuring complex learning outcomes and they can be more difficult to score than other types of questions, such as multiple-choice and true/false (Gronlund & Linn, 1990:145-146; Popham, 2003:88).

- **Binary-choice questions**

Binary-choice questions require choosing between two correct options such as: true/false, right/wrong, correct/incorrect and accurate/inaccurate (Popham, 2003:72). These questions can cover a wide range of content by providing many questions on a subject in a short time, but the main problem with this kind of question is that there is a tendency for guessing, which can result in a 50/50 percent chance of selecting the correct answer by chance (Airasian, 1997:157; Popham, 2003:72).

- **Matching exercises**

Matching exercises are easy to construct, easy to score and can measure large amounts of information in a relatively short time (Airasian, 1997:157; Popham, 2003:78). The disadvantages are that there is generally an overemphasis on memorisation, and less significant material is introduced in order to supply material for the matching (Gronlund & Linn, 1990: 159).

Conclusions drawn from the above discussion make it clear that traditional assessment tests have many shortcomings. These types of questions are very behaviouristic and

do not take into account new insights on learning (as discussed in section 2.2 *Theories underlying assessment practices*).

From the preceding discussion, it can be deduced that these tests are not necessarily a reliable barometer of what students really know. Students can have considerable knowledge on a particular subject though they may fail to answer specific questions correctly. A more serious problem is that students may answer questions correctly on a test, yet be unable to transfer their knowledge to other settings and situations. They may also be incapable of applying their knowledge to simple and complex everyday tasks.

### **2.5.2 Criticism of traditional assessment**

The vast condemnation of traditional assessment methods warrants attention in this chapter. First, traditional assessment has been berated for failing to show the essential connection between learning and assessment. Second, it does not respect the student or act in his best interests. Third, traditional assessment doesn't reflect the principles of cognitive thinking. Last, traditional assessment methods may narrow instruction and may restrict teacher freedom.

- **Misunderstanding of the connection between learning and assessment**

Perhaps the most valid criticism of traditional assessment is that instead of strengthening the connection between learning and assessment, in practice it actually illuminates the differences. This problem can be traced, the researcher believes, to the misunderstanding of the relationship between assessment and learning. This misunderstanding leads in many situations to the distinct separation of assessment from learning (Wiggins, 1993:3; Shepard, 2000:10). For example, teachers often reserve tests until after the summation of learning and utilise tests as a measurement of how students have performed, rather than how they are performing. Thus, test results have come to be valued as a means of assessing work already completed. They are rarely used to guide learning in progress.

Critics, such as Maeroff (1991:275) and Rust (2002:1) suggest that assessment must be viewed not as an end, but as a concrete means towards an achievable goal. Students need to be shown that they can successfully master learning tasks; they need to be guided towards accomplished performance. An initial step in making assessment more effective, according to Wiggins (1993:4) would be in changing the basic relationship that often exists between tester and student. He contends that once it is recognised that students are the primary “clients” of assessment, greater attention should be directed towards their needs, which is essentially to improve their performance.

Wiggins (1993:209), who views the assessor as a guide who helps the student prepare for real-life situations, faults traditional testing with relegating the student to an inactive role. He insists that, “the aim of education is to help the individual become a competent intellectual performer, not a passive ‘selector’ of orthodox and prefabricated answers”. He supports his statement by offering the analogy of a doctor who is considered an expert in his field, not just by virtue of the set of rules and propositions that he possesses, but rather because he acquires the skill to adapt those guidelines to the individual patient and his unique case history (Wiggins, 1993:231).

Stiggins (2002:761-763) takes Wiggins’ point of view even further by suggesting a plan for assessment that will not merely check learning, but rather advance learning. In order to promote assessment for learning teachers need to:

- Understand and clearly explain in advance the achievement objectives that their students are expected to reach
- Notify their students about the learning goals, in simple understandable terms from the onset of the learning process
- Become knowledgeable of assessment theories and principles and be able to devise exercises and scoring methods that accurately reflect student achievement
- Utilise classroom assessment to build students' confidence in themselves as learners and help them assume responsibility for their own learning
- Explain classroom assessment results into recurrent descriptive feedback (not

judgmental feedback) for students, affording them specific insight as to how to improve

- Continue to modify instruction based on the outcome of classroom assessment
- Involve students in regular self-assessment, with the same standards in tact so that students can witness their own development over time and feel they have control of their own success
- Vigorously involve students in communicating with their teacher and their families about their degree of achievement and improvement

Students who learn to supervise their own learning and make active decisions increase their success and obtain the foundation for lifetime learning (Stiggins, 2002:263).

- **Does not respect the student or act in his best interests**

Many alternative assessment practices are unfair to students since they often act against their interests. Multiple-choice questions, for example, which require students to choose between answers that are similar in appearance, are a form of deceit (Wiggins, 1993:117-118). They use distracters that are deliberately intended to lead students to incorrect choices. The word itself, ‘distracter,’ Wiggins (1993:117-118) warns, “hints at, intentional deception”. Another problem Wiggins (1993:130) cites is the “inherent ambiguities” in some questions that call for answers that require some degree of dialogue or discussion that is discouraged or forbidden. By not allowing students the opportunity to explain incorrect answers, we prevent them from showing any understanding that they might have (Wiggins, 1993:9). Brown (1990:7) and Cohen (1990:47) express similar criticism of traditional assessment, viewing it as an impediment to learning rather than an advantage.

An additional flaw of traditional assessment is that it is often based in secrecy (Wiggins, 1993:72, 96). Students are, at times, unaware of the type of questions they will be asked, sometimes they don’t know when they’ll be tested and the standards of judgment are frequently unclear. The decision regarding marks is also often arbitrary. Some test results are kept secret and the test paper is not released. It is unfair that, although a great deal often rests on test results, students are prevented from seeing

their graded paper and from being permitted to understand their errors. They are barred from any opportunity to learn from their mistakes and often are prevented from knowing where they have made mistakes. Tests that deny the use of resources give students the understanding that resources are not valuable tools, and that they can't be utilised in 'real-life' situations. This practice is detrimental in that it hinders their ability in the future to call upon resources, both technical and human to succeed in actual situations when we are expected to use resources (Wiggins, 1993:96). This implication that one is on his own, devoid of any advice or guidance not only does not reflect real life situations, but misleads students into believing that it is unnatural to seek help.

Additional problems with traditional assessment are enconced in questions that are unclear. There are usually no opportunities for students to clarify their understanding of questions that are inherently or deliberately vague. As a result, incorrect answers do not truly reflect understanding of the material tested. Another limitation of commercially available tests (multiple-choice tests), Burt and Keenan (1995:2) contend is that evaluators may not know whether an ESL student is having trouble with selected items because of "difficulties with reading, with the vocabulary or with the cultural notions underlying the test items". In other words, the student might very well know the material, but be at a disadvantage to providing correct or acceptable responses.

An additional complaint that is frequently directed at traditional assessment is that it doesn't really measure what students know and are able to do. "Understanding", Wiggins (1993:58) explains, "is best assessed by pursuing students' questions, not merely by noting their answers". While recognising the obvious convenience of traditional assessment procedures (which have been discussed in section 2.5 *Traditional assessment*, critics believe that traditional testing methods are a poor indicator of genuine ability or real knowledge. One can concede that traditional exams do test recall of isolated facts and the ability to answer quickly under pressure. Yet testing which consists primarily of isolated items does not test critical thinking and reasoning knowledge (Wiggins, 1993:47; Darling-Hammond, 1995:4-6).

- **Norm-referenced instead of criterion-referenced assessment**

Another disadvantage of standardised testing is that it is generally evaluated by norm-referenced assessment. As discussed in 2.3.3.6 *Norm-referenced assessment*, norm-referenced tests are interested in scoring students in relationship to other students for selection purposes. Not only do these tests primarily feature closed questions (questions that have only one answer), norm-referenced, standardised achievement tests give the impression that answers are always right or wrong (Barth & Mitchell, 1992:15). These types of questions disallow students the opportunity to develop individual strains of thought. In addition, they fail to encourage sophisticated thinking and an abstract approach to problem solving. A more serious problem is that norm-referenced assessment is intended to measure an individual student's performance in relation to others. It is fundamentally competitive and scoring is according to rank. Therefore, able students may be denied consideration or admittance to programs simply because there are a limited number of accepted applicants.

Criterion-referenced assessment, on the other hand, which does not often occur in traditional assessment, recognises the accomplishments of everyone who meets the given standards. The goal of criterion-referenced tests is to get a description of the specific skills and knowledge that each student can demonstrate (Linn & Gronlund, 2000:43). This type of assessment is suitable for performance and project evaluation, and can be useful in planning instruction. The obvious benefit of this type of assessment is that attention is given to individual student accomplishment.

- **Doesn't reflect the principles of cognitive thinking**

One of the most popular forms of traditional assessment is standardised testing (section 2.3.2.1 *Standardised testing*). It became popular in the 1970s when the disappointment in existing educational systems resulted in demands for accountability (Worthen, 1993:445). The objective was to search to identify the best students for acceptance and selection purposes. Brown (1990:8) shares the view that many of the characteristics of past assessment have resulted from the purpose to select, and not to

teach and learn. She (1990:5) believes that assessment should advance learning opportunities, rather than sort people into groups.

A disadvantage of traditional assessment is that as classroom assessment has gained importance, learning and assessment practices suitable for standardised testing are no longer relevant. In short, assessment needs are different today because learning goals are different. Torrance (1995:10) insists: "Assessment of simple rote responses and memorised information does not prepare students to assume positions and responsibilities in the 21<sup>st</sup> century." They need assessment that will develop problem solving skills and encourage creativity. Hanna and Dettmer (2004:35) agree that learners need to search for and actively process information.

Traditional assessment causes young children to possibly know certain things without really understanding them and without being able to make generalisations from specific examples to similar problems in different contexts (Torrance, 1995:1). He argues that more 'authentic' or realistic tasks than traditionally used call on students to develop higher-order skills and capabilities such as problem solving, investigation and analysis. These are the skills they should be learning (Torrance, 1995:2-3). Cohen (1990:47) shares the view that current assessment strategies are an inadequate measurement for the depth of education that a student receives especially as far as language is concerned. He contends that standards need to change over time, maintaining that the "whole language" approach to language learning leads to a wider interpretation of assessment than the simple assessment of grammar and spelling (Cohen, 1990:43).

Cohen suggests that, for example, instead of the traditional testing of the number of words spelled correctly on a twenty word test, it would be more appropriate to use the words in telling a tale or writing a one-page story. Cohen complains that although assessment focuses upon products of learning, it generally ignores the processes of learning. In addition, current assessment strategies neglect important areas, such as creative thinking and problem solving (Cohen, 1990:47). Keeping the above in mind, it is clear that traditional assessment procedures are lacking. The most detrimental aspect of traditional assessment is the disservice it does to the student, which will be discussed in the following section.

- **Narrows instruction and restricts teacher freedom**

A major criticism of traditional assessment is the “narrowing effect” that traditional tests can have on curriculum content and teaching methods.

The problem of standardised testing is manifold as it results in test-driven curriculum (Worthen, 1993:446). Since heavy weight is accredited to test scores, tests have a sizeable influence on what is taught. As a result, teachers are pressured to teach what is tested and in those exact formats. This is detrimental as it confines the curriculum and limits teachers’ freedom. An additional misuse of this type of testing is that it is considered as criteria for teacher and school evaluation. Teacher accountability has become the ability to teach well to the test. The use of standardised tests for statistical purposes focuses on test results and ignores the direct interests of the students.

The above-mentioned reasons clearly explain the dissatisfaction with traditional assessment methods. However, this dissatisfaction with traditional assessment has led to more effective and equitable means of assessment. The next section provides an identification and discussion of alternative assessment and its impact on learning.

## **2.6 ALTERNATIVE ASSESSMENT**

The shortcomings discussed in section 2.5, *Traditional assessment*, resulted in a search for methods that could overcome the flaws of traditional assessment. The pursuit led to an investigation of alternative means of assessment which could more effectively measure a student’s skills and knowledge while respecting and promoting the student’s best interests. Alternative assessment, which is also called performance-based assessment, direct assessment and authentic assessment, measures what students are able to do rather than what they know. Worthen (1993:445) refers to the proponents of alternative assessment who fault traditional assessment for sampling only “tiny snippets of student behavior” rather than looking at a bigger picture of a student’s work.

There are three main reasons for the increasing popularity of performance assessment

in the USA (Airasian, 1997:213):

- It is being proposed or mandated as a part of formal statewide assessment programs.
- Greater emphasis on pupils' problem-solving, higher level thinking and "real-world reasoning" skills warrant performance and assessment to demonstrate such outcomes.
- It is believed that performance assessments will give pupils who do poorly on selection-type tests the chance to demonstrate their achievement in alternate ways.

Although there is no single definition of alternative assessment, there is consensus regarding its main goal, which is considered to be "the gathering of evidence about how students approach, process, and complete real-life tasks in a particular domain" (Huerta-Macias, 1995:9). Worthen (1993:447) agrees that alternative assessment should give direct measurement of real performance on important tasks.

Unlike traditional testing, where students are evaluated on what they are able to recall and reproduce, alternative assessment "evaluates students on what they integrate and produce" (Huerta-Macias, 1995:9). Pierce and O'Malley (1992:2) similarly define alternative assessment as "any method of finding out what a student knows or can do that is intended to show growth and inform instruction and is not a standardised or traditional test". They state that specifically, alternate ways of assessing students consider variation in student needs, interests, and learning styles; and they try to integrate assessment and learning activities. In addition, they designate successful performance, stress positive traits, and provide formative rather than summative evaluation.

Macmillan (2001:196) explains that performance-based assessment or authentic assessment "is one in which the teacher observes and makes a judgment about the student's demonstration of a skill or competency in creating a product, construing a response, or making a presentation". According to Messick (1994:17), performance-based assessments transform educational assessments into worthwhile educational experiences that motivate and direct learning. Similarly, Oosterhof (1999:151) views

authentic assessment as tasks that require the actual use of a life skill beyond the instructional framework.

From the preceding discussion it becomes clear that there is a general agreement that alternative assessment implies the monitoring of the production or presentation of a finished product. The advent of this form of assessment has been met with interest and enthusiasm, as the obvious benefits to students and the learning process are monumental. Yet it is important to note that despite the ardor regarding alternative assessment, it has not been seen as the single panacea. There are some, such as Stiggins (1997:178) and Brandt (1992:35) who insist that in order to achieve maximum learning it is necessary to preserve traditional methods and integrate them with newer methods. They believe that traditional assessment has advantages that can strengthen student performance.

Alternative assessment provides students with many advantages (Stiggins, 2002:764), such as:

- They are more confident learners since they can see their own success.
- There is greater achievement since they are willing to take risks.
- Students understand what it means to be in charge of their own learning and make decisions that affect their learning.

### **2.6.1 Criteria for effective alternative assessment**

Despite slight variations, there is a general agreement about the factors that contribute to more effective assessment. These factors, which are lacking in traditional assessment, are significant features of alternative assessment. Apart from validity, reliability, transparency, etc. that are criterion for assessment; the following are also criteria for effective assessment.

#### **2.6.1.1 Clear instructions**

A fundamental feature of successful assessment is the provision of clear and explicit instructions. Assessment criteria must be explained clearly so that students can

understand them (Maeroff, 1991:278). The most significant element is the need for clear decisive instructions and guidelines for performance tasks. It is apparent to the researcher that only when tasks are carefully outlined and explained can students have the direction they need to complete a task successfully. Good assessment should also consider the learning styles, language abilities, cultural and educational backgrounds, and grade levels of the student.

Clear instructions are essential for effective teaching. Teachers need to clearly explain what students are expected to know, understand and be able to do as a result of the instructions (McTighe, 1997:6-7). They need to provide evaluative criteria and models of work that exemplify different stages of quality so that students have examples to emulate. Rubrics that are clear and precise can be a valuable guide to student performance (McTighe & O'Connor, 2005:12).

Students demonstrate their learning in different ways and so they should be offered a choice of tasks that can be adapted to accommodate their individual needs. They should be allowed to utilise their own strengths and be permitted to choose suitable options for displaying their knowledge, skills and understanding (McTighe & O'Connor, 2005:14-15). For example, students with graphic skills could be directed towards implementing these skills in their projects, while students with musical skills could wield their particular ability. When teachers allow and encourage students to take advantage of their talents, they are able to showcase their strengths, rather than weaknesses. Teachers learn about their students' weaknesses when they observe them engaged in a variety of tasks.

In addition, in order to carry out a performance task, teachers need to make sure their students have the prerequisites or skills necessary to work efficiently (Dakwar, 2003:56). Students might need review in using appropriate registers, grammatical structures and writing strategies. In addition, they might need to be taught various skills, such as listening skills, critical thinking skills and group process skills (Herman, 1992).

### **2.6.1.2 Continuous assessment and opportunities for improvement**

Another important factor of assessment is monitoring the individual student's growth over time, rather than at the culmination of an activity (Tannenbaum, 1996:1; Baker, O'Neil & Linn, 1991:38). Since real learning is an ongoing process it is important for students to produce, reproduce and present their work after corrections and modifications have been done. Students should be shown that genuine performance or production is really work in progress. They need to see the value in reviewing and refining their work. The psychological advantage of being permitted to adjust an assignment is immeasurable. The benefits are twofold. First, students understand that real experiences offer second chances to correct mistakes and misunderstandings. Certainly, students will feel less threatened when they know that they will be permitted another chance. Second, teachers make it more possible for their students to reach educational objectives, which is the real purpose of education.

### **2.6.1.3 Feasibility**

Assessment must be feasible, which means that it must be the finest assessment of a learner's knowledge and ability within the perimeters of time and resources (Kilfoil & Van der Walt, 1997:293). To be practical, assessment costs must be kept at acceptable levels and assessment methods must be practical within the limits of available time.

### **2.6.1.4 Authenticity**

In order to maximise learning, alternative assessment tasks should be authentic in order to generate interest and present a genuine challenge to the students. According to Linn et al., (1991:20) tasks should get students to deal with meaningful situations that proffer valuable educational experience, which can be transferred in some degree to other real-life challenges. When students see that they can apply what they learn in class to actual situations they encounter, they will be anxious to invest their energy in perfecting these skills.

Therefore, attention to content is an important criterion for judging any assessment task. Linn et al., (1991:20) advise that the content of an assessment task must be

worthy of both students' and raters' time and effort. Students are interested in material they consider to be relevant to their lives and experiences. Teachers are willing to spend more time in assigning and assessing tasks that they consider to be important and engaging.

In addition, Gardner, who believed strongly that the goal of education is to have people who are able to do things in the world, advocated the need to teach pupils to do those things and to assess these performances accordingly (Gardner, 1994). He offers harsh criticism of educational systems that for varied reasons (ideological or financial) teach and assess all students in a similar way and believe that they have treated everyone fairly. Gardner in an interview with Brandt insists that the findings of cognitive research over the past twenty to thirty years reveal that students “do not understand, in the most basic sense of that term”. He believes they lack the capacity to take knowledge learned in one setting and apply it appropriately in a different setting, which is evidence of genuine understanding (Brandt, 1993:4).

#### **2.6.1.5 Student-centered**

Eisner (1991a:14), who contends that maximal learning occurs when classroom activities are created with student interest in mind, sees something unnatural in directing students to pursue tasks that have been created for them. He finds fault with the assessment practices that are most frequently used. The message to conform that they transmit diminishes what he calls students' “sense of personal signature” (Eisner, 1991a:15-16). Although he recognises that an identical outcome is desirable in certain tasks, such as spelling a list of words correctly, students should be given expressive activities (creating a story, painting a landscape, choreographing a dance or composing a piece of music) that will yield diverse outcomes. Eisner (1991a:17) suggests that a good school “does not diminish individual differences, it expands them”.

An important objective of education should be to provide children with chances to learn and formulate their own goals, and design their own strategies for achieving them (Eisner, 1991a:14-16). Since assessment is an educational tool that should benefit students, all elements of assessment ought to work jointly to facilitate,

evaluate and further guide learning. Since assessment directly involves and affects students, it must be the focus of the assessment process. However, projects and assessment are most often designed and controlled by the teacher.

#### **2.6.1.6 Fairness**

The criterion of fairness needs to be applied to any assessment (Sieborger & Macintosh, 2004:11). Whether assessment is traditional or alternative, the potential for bias can exist. The implementation of alternative assessment tasks raises several issues about fairness. Students naturally have differences in familiarity, exposure, and motivation regarding the tasks (Linn et al., 1991:18). Therefore, they will not relate to their work in the same way. Since not all students have equal access to resources, it is the researcher's opinion that teachers have to provide necessary resources for all students who need them. For example, if computer technology is not available for all students, or if some students have limited or no access to libraries and other research sources, teachers need to bring relevant research material into the classroom so that it is available to all the students.

Teachers must also be sensitive to the fact that certain subject material, which might be interesting to some students, could be insulting to others. For that reason, procedures for identifying offensive materials or materials that present a particular challenge to specific student groups need to be implemented (Linn et al., 1991:18).

Fairness does not imply that there should be lower expectations for different learners. It suggests that it is possible for students to produce different types of evidence that they have reached expectations. Student performance need not be identical in order for the grading system to be fair (Loyd & Loyd, 1995:483).

Gardner (1993:176) refers to "intelligence-fairness" which he believes can be achieved by assessing learners on the use of multiple intelligences rather than on logical/mathematical and linguistic abilities.

The following five factors should be taken into account when trying to establish fairness (Sieborger & Macintosh, 2004:13):

1. The importance, length, size or weight of the assessment - does it match the amount of work that has been done?
2. The choice of the assessment technique - does it match the way in which the learners have been taught?
3. The instructions and/or questions given to the learner - do all learners understand what they are expected to do or answer?
4. The method of administering the assessment - are the conditions appropriate; is there enough time; do learners have access to the resources that they need?
5. The method of marking - is it as objective as possible? Should the learners know how their work will be marked?

Clearly, in order to provide fair assessment the diverse needs of the students have to be considered. Teachers must understand and recognise their students' differences and strengths in order to assess them fairly.

#### **2.6.1.7 Improvement of learning**

Alternative assessment can be a valuable tool for improving learning. Firstly, the creation of projects and the delivery of performances demand that students acquire and hone a variety of important skills. For example, if debates, discussions, essays and problem solving are incorporated into the testing system, children will spend time practicing these important activities (Resnick & Resnick, 1992:59). Different areas of performance can be incorporated into classroom assessment activities, such as reciting poetry, constructing original stories, designing book covers, conducting surveys, etc. This type of assessment lends itself to innumerable diverse activities, which can whet student appetite and spur interest.

Secondly, if principles of alternative assessment are incorporated into the curriculum, they could bring about significant improvements and advances in education. For example, Torrance (1995:2,150) suggests that changes in assessment might actually precede curriculum changes, and could be the driving force behind beneficial reforms in the curriculum. He explains that by providing more complex and demanding assessment tasks, educators can broaden the curriculum and raise the standards of teaching.

Thirdly, by regarding assessment as an interactive process both teacher and student become involved in overseeing student performance. The importance of student involvement has been discussed in section 2.5.2 *Criticism of traditional assessment*. Therefore, students should be provided with opportunities before and after units of instruction for self-assessment so they can assess their own performance (Spolsky, 1992:38; Sieborger & Macintosh, 2004:8). Students could have some choice in deciding which tasks will be carried out and which criterion must be met (Sieborger & Macintosh, 2004:8). For example, for a book task they could decide whether to write a letter to the author, compose a new ending to the story or research the historical period that provides the setting of the story. Students could also decide which criteria are important for a specific task, which would lead to discussions of various language skills, such as syntax, content development and vocabulary.

When teachers encourage students to take part in decisions concerning their own learning, they feel a sense of worth and importance. Moreover, to enable them to make educated decisions, students need to use skills and knowledge in order to think and discriminate, and make choices. Students who are involved in all aspects of assessment have additional interest and greater stake in their performance.

Finally, the emphasis in alternative assessment should be on tasks that incorporate higher-order or complex skills (Baker et al., 1991:38). Most usually these tasks involve the design of a project or performance. When students are engaged in planning, preparing and executing a project they call on skills and knowledge they have acquired over time. Moreover, they learn additional skills and venture into new areas of study and experience.

#### **2.6.1.8 Takes place in the classroom**

There is an additional problem that is inherent with assessment that is created and controlled outside the classroom. Hamp-Lyons and Condon (1999:26) fault this type of assessment for having no particular relation to what teachers are teaching in the classroom. He suggests that performance assessment is a far better means of evaluating student work as it presents fuller information about what students can do, and gives students the chance to actually do more than the test requires. Similarly, in

his condemnation of the “unacceptable state of classroom assessment”, Stiggins (2001:5) criticises assessment for being disconnected from daily instruction. He chastises teachers for allowing student assessment to be controlled by measurement experts who are disassociated from classroom instruction.

Stiggins insists that classroom teachers need to take charge of classroom assessment and moreover they need to become familiar with a variety of assessment techniques, which they will use in the classroom. He emphasises the need to convert the classroom into a safe haven, where students can feel comfortable taking risks and attempting the unknown. He suggests that “maximising students’ belief that they can learn should replace the once held view that maximising anxiety will maximise learning” (Stiggins, 2001:5). The researcher supports Stiggins’ contention that assessment should be used to build student confidence, not to inhibit or retard. He also acknowledges the role that students should play in assessment by record keeping, and self-reflections, which will focus their attention on progress (Stiggins, 2001:12).

### **2.6.2 Alternative assessment methods**

Whereas traditional assessment is identified primarily with written tests, alternative forms of assessment are varied and numerous. They all, however, require students to show what they are able to do with their knowledge and skills. In addition, they are usually evaluated at different stages of progress and are assessed according to clearly designated criteria in the form of checklists and rubrics.

The advantage of alternative assessment is that in contrast to the “depersonalised appraisal” of traditional assessment, there is “a refreshing degree of intimacy” achieved through “thumbing through a portfolio with a student or watching a student perform a task” (Maeroff, 1991:281). It is clear that alternative assessment has the potential to enrich the relationship between teachers and students, and to enhance learning.

The most popular examples of alternative assessment are: performance assessment, self-assessment and peer and group assessment. These activities are promoted by advocates of alternative assessment.

### **2.6.2.1 Performance assessment**

McMillan (2001:196) defines performance assessment as “one in which the teacher observes and makes a judgment about the student’s demonstration of a skill or competency in creating a product, construing a response, or making a presentation”. Messick (1994:17) expands on McMillan’s definition by stating: “Performance-based assessments are thought to transform educational assessments into worthwhile educational experiences serving to motivate and direct learning because students are required to demonstrate their understanding”.

Educators, like Rust (2002:3) agree that the type of assessment chosen should be connected to the desired learning outcomes. They caution that it is important to consider the purpose, validity and relevance of an assessment task. In addition, there should be a variety of assessment activities so that students can choose activities that they feel comfortable with.

The types of performance assessment used most often in classrooms are: performance-based tasks, portfolios, journals and diaries and projects.

#### **(a) Performance-based tasks**

Performance-based tasks, a popular form of performance assessment, are tasks that require students to employ the knowledge and skills they have acquired by creating a product or delivering a performance. Hibbard (1996:5) define performance-based assessment as a demonstration of a set of strategies for the use of knowledge, skills and work practices through the performance of tasks that are significant and appealing to the student.

These tasks, though varied in nature, are based on general principles and designs of evaluation (Gronlund & Linn, 1990:6-7).

These include:

- a) Clearly specifying what is to be evaluated (clearly specifying the intended learning outcomes before selecting the achievement measures to use)

- b) Selecting an evaluation method in terms of its relevance to the characteristics or performance to be measured
- c) Employing a variety of evaluation techniques
- d) Being aware of the limitations of evaluation techniques such as the chance factors; guessing, subjective scoring on essays, misinterpretation of test results etc.
- e) Understanding that evaluation is a means to an end, not an end in itself.

Examples of performance tasks would be to act out a scene from a book, write and deliver a defence for a character in a story or create a time line of the events in an historical novel.

Herman, Aschbacher and Winters (1992:25-26) suggest five questions that educators should ask themselves when evaluating a performance task.

1. What important cognitive skills or attributes do I want my students to develop? (e.g. to communicate effectively in writing; to analyse issues using primary sources and reference materials).
2. What social and affective skills or attributes do I want my students to develop? (e.g. to work independently, to work cooperatively with others, to have confidence in their abilities, to be conscientious).
3. What metacognitive skills do I want my students to develop? (e.g. to reflect on the writing process they use; to evaluate the effectiveness of their research strategies; to review their progress over time).
4. What types of problems do I want them to be able to solve? (e.g. to undertake research; to solve problems which have no single, correct answer).
5. What concepts and principles do I want my students to be able to apply? (e.g. to understand cause-and-effect relationships).

Considering these questions is imperative for teachers prior to initiating any performance-based task. Teachers need to direct their activities with specific objectives in mind. These points need to be decided upon before students begin their activity. The performance task, when completed, should give evidence that the intended goals have been met. Carefully planned tasks and attention to desired learning outcomes will benefit both students and teachers.

Some experts insist on the implementation of performance-based instruction in the classroom on a regular basis. McTighe (1997:6) suggests principles of performance-based instruction that teachers should incorporate, such as clarifying the expectations of what students are expected to know, understand and be able to do as a result of instruction. The performance tasks chosen should represent an authentic exercise that the students could be expected to do in the real world, for example utilising the skills and knowledge they have acquired to produce a product or present a performance for a specific audience.

It is extremely important that criteria whether in the form of a rubric, a rating scale or a performance list are provided to guide students in the development of their products or performances. McTighe suggests that in addition to providing scoring criteria students should be given direction in applying these criteria to their own work by providing them with demonstrations or other models of the kind of work they are expected to accomplish. McTighe (1997:6) is not the first to suggest that novices need models as examples on which to base their performance.

Wiggins (1993:49-53) argues that an authentic assessment system has to be based on known, clear nonarbitrary standards, as students cannot be expected to master complex tasks if the criteria and standards are not revealed. He cites as examples the use of sheets of music, scripts, database rules, and game strategies, which are acceptable criteria, as opposed to test material, which is often kept secret. Certainly, when students are presented with clear criteria and examples of work that has successfully achieved the targeted objectives they are best equipped to imitate desired standards. Educators need to provide explicit understandable directives to guide student performance. They are remiss if they fail to do so.

### **(b) Portfolio**

A widely used learning instrument is the portfolio, which is a “purposeful collection of students’ work that demonstrates to students and others their effort, progress, and achievements in given areas” (Genesee & Upshur, 1996:99; Smith, 2002:20; Arter, 1995). The portfolio has advantages for both students and teachers. Since the portfolio is confirmation to learners of their attempt, growth and accomplishment it is

very appropriate for self-assessment as well as teacher assessment. An additional benefit of portfolios is that they “encourage ownership, pride, and a sense of self-esteem”(Frazier & Paulson, 1992:64).

In order to assure the success of a portfolio, Vavrus (1990:50-53) proposes decisions that should be considered before a portfolio is assigned:

1. It should be decided whether the structure is physical (an actual arrangement of documents that are used to show student progress) or conceptual (containing underlying goals for student learning).
2. A decision should be made regarding what goes into the portfolio and who the audience is.
3. Decisions regarding portfolio submission times and the explanations regarding the selection of material should be made with the students.
4. Evaluation standards should be clearly established prior to assigning a portfolio (standards can measure individual growth over time rather than comparisons among different students).
5. A decision should be made regarding what happens to the portfolio after the semester or course of study is over (should they remain with the student or be passed on?).

Arter (1995) suggests several elements that should be considered when assigning portfolios:

1. Teachers should have a clear idea of the student skills that they want to focus on.
2. Students should be involved in choosing what is included in the portfolio.
3. There should be self-reflection through which students communicate what they think about their work, their learning and themselves.

Smith (2002:21) suggests a framework for portfolios that would include:

- Introduction – The student will introduce him or herself and explain what will be presented in the portfolio
- Compulsory part – Core tasks of the course. Reflection on process and product
- Personal choices – Individual choice from a proposed list of tasks. Reflection

on choice, process, and product

- Reflections – Reflection on the learning process
- Conclusions – Message to the reader about the content

Genesee and Upshur (1996:99) recommend that in order to maximise assessment, students could be asked to answer guiding questions, such as: a) To what degree did you achieve your learning goals? b) To what extent did you improve your reading? c) Which problems did you encounter while reading and how did you solve those problems? d) How has your vocabulary improved? List the new words you have learned from your portfolio. e) To what degree has your confidence in English improved?

Clearly, there is flexibility regarding the contents of a portfolio. According to the researcher, the greatest benefits can be gleaned when students decide what to include in their portfolio. Their reflections can be an effective tool in helping students evaluate their own learning progress. In addition, the reflections can be valuable in helping guide teachers towards effective instruction.

### **(c) Journals and diaries**

Student journals and diaries are additional examples of alternative assessment writing activities. Generally these are less structured than portfolios and the main aim is to maintain a steady stream of communication between the student and the teacher. The premise behind journal and diary writing is that if students are encouraged to write in the target language, they will be more interested in expressing themselves clearly and correctly. In order to facilitate written communication in this way, it is important to stress the communication of ideas and feelings, rather than focus on attention to correct spelling and grammar (Sheppard & Stoller, 1995; Ur, 1996:168).

### **(d) Projects**

A project is “an extended investigation carried out by an individual student into a

topic agreed on by student and assessor” (Rust, 2002:3). The advantage to project work, I believe, is that the student assumes responsibility for the choice of subject and method of presentation. Project work offers students the opportunity to work independently or in groups for group projects. Projects are a way of empowering students by giving them the responsibility of choosing a topic and managing the development of their work (Orkin, 2004:37). They have the chance to follow through an assignment from the planning stages to fruition and monitor their progress along the way. The researcher thinks that projects provide students with an invaluable learning experience. Since they are active participants in every phase of their project from inception to completion and reflection, students understand that they themselves are the decisive factor in their ultimate success or failure.

A more extensive discussion of performance-based tasks is presented in Chapter 4, which deals with development in foreign language teaching and learning.

#### **2.6.2.2 Self-assessment**

Self-assessment takes place when learners assess their own work. Learners who recognise their own progress will be more motivated to stay on task. They will see the value of what they are learning. Harris (1997:2) explains that by appealing to the affective components (we have already seen the importance of the cognitive factor) students might be receptive to spending longer time on language study. Ultimately, as Murphy (1995:14) maintains, the most important evaluation is what students think of themselves.

The obvious psychological benefit is that self-assessment can ease student discomfort (Harris, 1997:19). Like many experts, Harris (1997:19) believes that assessment by nature of its traditional purposes (i.e. to rank, to identify weaknesses) is a source of fear and anxiety. Learners, he believes can be helped to perceive their own progress and encourages them to see the value of what they are learning. He views self-assessment as a practical tool that should be integrated into every classroom activity. Harris (1997:13) believes that self-assessment can motivate learners into becoming more active in their learning process. It can help them locate their own strengths and

weaknesses, and perhaps, most important of all, it can help them realise that the ultimate responsibility for learning is theirs.

An important facet of self-assessment is that it engages the students in addressing the assessment criteria for themselves (Rust, 2002:4). Evidence indicates that students acquire greater insight into what is actually required when they attempt to employ criteria or use model answers to mark their own work (Rust, 2002:4). Self-assessment is a worthy tool for placing students in the center of assessment where they belong. Self-assessment and reflection help learners gain a fuller understanding (Wiggins & McTighe, 2006:27) and make them more aware of their own strengths (Sadler & Good, 2006:2).

Self-assessment is not a luxury, as some would believe, but rather a vital element of formative assessment (Black & Wiliam, 1998:143). It helps students improve significantly and results in increased learning (Sadler & Good, 2006:1).

### **2.6.2.3 Peer assessment and group assessment**

Peer assessment is similar to self-assessment, which places the responsibility of assessment on the student. In peer assessment, however, fellow classmates use established criteria to assess each other. This method is advantageous as it focuses attention on criteria that is important for their own work as well as that of the other students whom they are assessing (Rust, 2002:4). An added benefit is that students become capable of expressing themselves clearly to others (Plavin, 1998:44). Whereas students might feel inhibited about having their work assessed by their teachers, they may feel more comfortable with peer assessment (Plavin, 1998:44).

Tudor (1996:182) states that peer assessment is a practical form of learner training. Since learners feel less stress when they identify the same mistakes they make in others, they become more receptive to achieving shared goals. Assinder (1991:218-28) agrees that when students prepare and deliver tasks for each other they benefit for several reasons. They have increased motivation, they have the opportunity for authentic communication, they can glean in-depth understanding, and they experience

feelings of commitment and confidence. As a result, practice becomes more meaningful and accurate.

Another form of peer assessment is group assessment. There are many advantages of group work and group assessment. Students who work together and assess each other can share information, and contribute ideas. They can benefit by listening to others, learning to wait their turn to speak, and feeling a responsibility to the group to complete assignments on time, and to share valid and productive criticism (Airasian, 1997:218).

Since learning is by nature a social activity, effective participation in class groups facilitates learning and teaches learners to work cooperatively.

## **2.7 ASSESSMENT IN SECOND AND FOREIGN LANGUAGE TEACHING**

Until recently, assessment in EFL/ESL classes was governed by summative evaluation, the focus being on linguistic accuracy and distinct language features, rather than communicative competence (Rea-Dickins & Rixon, 1997:151). In recent years, however, there has been a greater concentration on the importance of communicative proficiency in language study. This is the result of increased dependency on language as a communicative tool. Therefore, communicative teaching methodology calls for a considerate emphasis on learning to use language in authentic settings and the evaluation of language proficiency in realistic terms. As a result there has been a dramatic change in language instruction and assessment. As instruction methods change in order to incorporate new standards and principles, assessment strategies have been adjusted accordingly. For example, alternative assessment, in the form of self-assessment, peer assessment, learning diaries and portfolios are good assessment methods as they are more personal and allow for greater student freedom.

### **2.7.1 Benefits for language teachers**

Alternative assessment provides many benefits for language teachers. According to Stiggins (2002:764) these benefits are:

- Their students are more motivated to learn.
- Teachers make better instructional decisions because they understand assessment better.
- Teachers save time since they are able to develop and use classroom assessments more efficiently.

Popham (2003:18) also recognises the benefits that alternative assessment offers teachers. He points out that with alternative assessment:

- Teachers must identify the knowledge and skills that students need to know and this helps them plan their lessons and teach better.
- Teachers need to have a clear understanding of assessment objectives which helps them explain more clearly.
- Teachers provide immediate feedback which is an important element of productive instruction.

Steiner (2000:10) agrees that standards benefit teachers since they help teachers design curriculum, instruction and assessment. As a result, pupils are better able to understand what is expected of them, which facilitates learning.

The shift to alternative assessment encourages language teachers to nurture a vibrant learning environment where their students can interact and cope with new language (Hawkins, 2004:6.) It challenges teachers to develop critical and reflective thinking practices in their teaching and forces them to engage themselves in critical thinking (Hawkins, 2004:6).

The changes in the approach to assessment that have been adopted around the world have had a tremendous influence on the teaching and assessment of English as a foreign language in Israel. This influence and the monumental changes in language teaching, in general, and particularly in Israel will be discussed in depth in Chapter 4.

## **2.8 RELATIONSHIP BETWEEN STANDARDS AND ASSESSMENT**

Assessment in general and assessment in language teaching in particular has been

greatly influenced by the standards movement. The attention to standards has thrust assessment into a more prominent position in the learning process. Whereas the planning of a lesson was the initial stage, followed by student evaluation, there is now a reshuffling of priorities. In current practice it is wiser to first establish the learning goal, consider the evidence that will show that students have mastered goals and finally design the lesson around these factors. This practice, referred to as *backward design* (McTighe & Thomas, 2003:52-53) as well as other practices spurred by the standards movement is discussed in detail in Chapter 3, *The implication of standards in education and its impact on assessment practices*.

## **2.9 SUMMARY**

This chapter has intended to provide a comprehensive look at the role of assessment in learning. Traditional assessment methods, which were popular for a considerable period of time, fail to meet the needs of contemporary education. Clearly, the cognitive revolution illuminated the concept that the mind has a vital and active role in understanding and construction. This has resulted in the implementation of alternative methods of assessing knowledge and skills. The important contribution of alternative assessment is that a greater emphasis is placed on tasks that require problem-solving, comprehension, critical thinking, and reasoning. In addition, alternative assessment has brought the student to the center of the learning process, which has a monumental effect on how students react and relate to education. Students need the opportunity to direct and manage their own learning. Assessment activities, such as portfolio reflection, can help them develop an awareness of what they have learned. Self-assessment and peer and group assessment encourage students to take responsibility for their own academic success.

It is clear that many factors have influenced assessment procedures over the years. The standards-based movement has had, perhaps, the greatest influence on assessment in general and on assessment in language teaching in particular. The reasons for the advent and development of standards-based education and its subsequent influence on teaching, learning and assessment are discussed in the next chapter.

## CHAPTER 3

### THE IMPLEMENTATION OF STANDARDS IN EDUCATION AND ITS IMPACT ON ASSESSMENT PRACTICES

#### 3.1 INTRODUCTION

The field of education has gone through a major metamorphosis in the past several decades. The monumental changes that have taken place can be traced to four significant influences. Firstly, dissatisfaction with existing school systems resulted in modifications in teaching practices. The school reform in the United States in the 1980s was, for example, a direct response to the feelings of disappointment and displeasure with the existing educational system that was severely criticised for failing to produce a sufficiently educated population (Marzano & Kendall, 1998). Secondly, the role of education was seen in a more contemporary light. As it became clear that education was the major vehicle that could prepare young people for the increasing challenges of modern life in the 21<sup>st</sup> century, radical adjustments in education needed to be made (Lewis, 1995:746).

Thirdly, as more divergent views concerning knowledge acquisition began to emerge, educators were pressured to respond to the newly discovered patterns of learning. For example, development in the fields of understanding and a fascination with the intricacy of the human mind became closely entwined with investigations into the relationship between thought processes and the mind's response to educational stimuli. This resulted in widespread changes in educational practices, in which greater efforts have been invested in teaching methods that could best assure academic success. An example of this can be seen in the adaptation of various alternative instructional and assessment practices, which abandoned behaviorist philosophies for the more progressive theories of constructivism, cognitivism and multiple-intelligences (that were previously discussed in sections 2.2.1-2.2.4).

Fourthly, as English continued to gain strength as a lingua franca of the world, a major influence on education reform, particularly in the area of English language study, were changes in the teaching of English as a second language or foreign

language. As English gained more and more prominence as a global language, which directed the world of modern technology (especially for Internet usage) and development, it became strikingly clear that students would need to master greater language proficiency than in the past in order to function effectively (Sifakis & Sougari, 2003:64). Therefore, better and more efficient methods of teaching the language were required. Moreover, greater attention needed to be directed at training students to use language in real life situations, where mastery of correct language usage is vital.

A major advancement in the quest for more effective instruction was the creation of educational standards. Educators in the United States, like Resnick and Nolan (1995:8) and Scherer (2001:14) came to believe that by identifying, constructing and adapting educational standards they could provide equal learning opportunities for all students and facilitate and guarantee successful academic achievement. They based their belief on the fact that countries, such as Japan, France, the Netherlands and Germany, which were known for their exceptional students, have national curriculums based on standards (Resnick & Nolan, 1995:6). Other countries around the world, such as the United Kingdom, Canada, Italy, Australia and later Israel, became anxious to implement standards hoping to duplicate the academic achievement of more successful countries.

Standards (also referred to as assessment standards) have probably had the most significant influence on school reform. The influence is evidenced strongly in the United States. Lewis (1995:745) maintains that “whether lauded as a sign of progress or scorned as anathema”, the standards movement is one of the most talked about issues in school reform. Smith (2004:83) acknowledges that the influence of a standards-based education plan has been accelerating all over the world.

The standards movement initiated a customised approach to education, which resulted in a significant impact on assessment (Reigeluth, 1997:204-206). Multiple levels of achievement, recognised by the various assessment standards, necessitated a variety of assessment activities. Since standards require students to perform a range of proficiency, it is necessary to provide them with a variety of assessments to measure

the learning of these competencies (Kluth & Straut, 2001:45). The single standardised approach to assessment became no longer valid.

In addition, the shift to communicative language teaching necessitates alternative assessment. Students need to be presented with assessment activities that encourage the use of authentic language in real-life situations. Traditional pen and pencil assessments consisting of multiple-choice and matching questions do not require students to demonstrate language competency or proficiency. Alternative assessment activities necessitate authentic language activity and interaction.

The discussion of standards in this chapter concentrates on the educational standards that provide the base for school curricula. These standards have been created and implemented as a foundation for educational systems around the world on local, state, national and international levels. Although the precise number and description of these standards differ, there is consensus that a unified body of standards can improve teaching, learning and assessment in schools. The race to adopt standards in countries around the world influenced the adoption of standard-based education in Israel.

### **3.2 DEFINITION OF STANDARDS**

Although there has been sufficient agreement to adopt national standards in the USA, there has not been a unanimous understanding of what the term “standards” mean. As already mentioned, the Collins Cobuild English Language Dictionary (1992:1421) defines standards as “a level of quality or achievement, especially which is thought to be acceptable. It is something used to measure or estimate the quality or degree of something, for example, how good a piece of work is”.

There are numerous additional explanations of standards. O’Neil (1993:6), for example, explains that standards are:

- What students are required to know and be able to do
- The vital basic knowledge in a particular subject area
- A passing mark on an evaluation
- An exemplary demonstration worthy of copying

Smith (2004:83) explains that standards are agreements about what is essential in education; she points at a general nucleus of knowledge and skills which learners are expected to attain. Wiggins (1991:19) defines standards as instructive, explicit examples of excellence on assignments we value. Most important, he sees standards, not as theoretical ambitions or unattainable aspirations, but rather as precise and directing depictions of laudable objectives (Wiggins, 1991:21). It is clear from this litany of explanations that standards imply not only the satisfactory accomplishment of essential goals, but also imply reaching a level worthy of imitation.

Educators such as Sabers and Sabers (1996:19-21) and Wiggins (1991:23) firmly insist that standards in education refer specifically to high expectations, which advance high levels of student achievement. Therefore, it is clear that reaching standards infers not merely to completing a given objective, but rather to reaching significant academic heights. The question naturally arises as to what educators can do to ensure that their students aspire to and attain high scholastic goals. An excellent example of efforts in this direction can be seen in *The Professional Standards for Teaching Mathematics* (1991), a document developed by the Working Groups of the Commission on Teaching Standards for School Mathematics of the National Council of Teachers of Mathematics (NCTM) in the United States. This document recognises the importance of empowering students by having them engage in mathematical reasoning, conjecturing, inventing and problem solving. It promotes a shift away from isolated mathematical exercises to an approach that connects mathematical ideas and applications (Frye, 1991).

These guidelines set by the NCTM became a blueprint for standards documents drawn up around the world, including the United States (Scherer, 2001:14). Yet once an agreement is accepted regarding the meaning of standards, the first step in the development of standards must be to understand the reason for incorporating standards into a system. Standards should be included in an educational network only when there is a clear, direct understanding of what standards hope to achieve.

From the above explanations the following definition of standards could be concluded. Standards are the optimal level of accomplishment that one is expected to reach in a particular area of study. In this thesis, this definition will be used.

### **3.3 THE RATIONALE BEHIND STANDARDS**

There are several reasons why standards have ignited considerable enthusiasm. The most important reason for standards is the extensive criticism of the educational system (Smith, 2004:83). Ranking students according to districts, states and even countries has brought attention to how students measure up on a large scale. Pressure arises when student populations fail to measure up successfully to students in other places. When the competition takes on a worldwide audience the need arises to achieve international ranking. When schools or students fail to compete successfully, educational bodies search for the cause. The next step is to insist on some kind of change or reform that can alter the results.

The implementation of standards has been embraced as a solution to the problem. Gratz (2000:681) questions whether reforms that focus on “world class” standards will succeed. He fears that standards, which are adopted for political, rather than educational reasons will not be implemented well. Furthermore, he predicts that if standards-based education fails, teachers and students will be unjustly blamed.

#### **3.3.1 Desire for greater academic achievement**

Smith (2004:83) points out that although the wide generality of standards is advantageous as it represents the kinds of achievements that we want all students to attain, it is also detrimental as students’ motivation, ability and personalities are not all the same. In addition, there is the plight of minority students who will find it difficult or impossible to fill the gaps in their education. Therefore, it seems futile to assume or expect that diverse students can reach identical standards.

Gratz (2000:681-2), however, stresses the positive result that can be achieved by demanding more of students. He blames low achievement on America’s inclination to excuse the low achievement of minority students, which results in a widening of the existing gaps. Gratz (2000:686-687) who views standards as a conduit for eliminating the inequality between high and low achieving students, maintains that by offering improved standards for everyone the level of education will be elevated. Tucker and Coddling (1998) agree that it is detrimental to accept lower standards. They insist that

other countries are more successful academically since they expect students to meet educational challenges and, consequently, they do.

### **3.3.2. Accountability**

Standards have been introduced to meet the demands for accountability (Smith, 2004:83). In other words, when students fail to perform successfully, the search begins for those responsible for the failure. In most cases the term “accountability” is a term used by politicians and heads of institutions (Smith & Fey, 2000:334). It signifies the requirement of one group to provide justification to another group in return for trust and dispensation (Sachs, 2003:177). Accountability in education means that schools are expected to meet specific educational goals and when they fail to do that they are berated for not achieving the necessary expectations. Consequently, reforms are devised to remedy an ailing situation.

Standards that are created as a result of accountability are not effective standards as they are in this case set by goals determined outside the classroom. These standards are developed in a “top-down procedure” usually by politicians who seek a quick answer to low test scores, and are ineffective (Smith, 2004:83). In order to develop truly operational standards, teachers and educators in the field need to be part of the process. They are in the best position to understand and establish the scholastic objectives of their students. Furthermore, they are most equipped to guide their students to the fruition of these goals.

### **3.3.3 Guidelines for teachers**

Standards operate as guidelines for teachers and learners in executing their assignments (Smith, 2004:83). They help teachers plan, teach and assess by specifying the educational objectives that must be met. Students benefit from knowing what is important to learn. In addition, they know what they are expected to accomplish. Once standards are clearly set, assessment becomes more viable as a measurement of targeted knowledge and skills.

Standards are also an invaluable tool in helping to direct teachers towards designing a curriculum and instruction that will afford opportunities and establish high expectations for all students in the 21<sup>st</sup> century (Kluth & Straut, 2001:46). Standards can help to merge the official curriculum with what is taught, what is tested and what is actually learned and thereby maximise learning (Sandrock, 1997:10-11).

### **3.4 PRINCIPLES UNDERLYING STANDARDS**

Since considerable weight has been given to standards-based education and assessment it is essential to discuss the principles surrounding standards. It is important to understand what considerations have to be taken to ensure that quality standards are developed. Sabers and Sabers (1996:21) complain that while most of the discussions about standards have focused on their importance, not sufficient effort and time has been invested in their content and to the consequence of their implementation.

Standards can be used in different ways and with different objectives in mind. Wiggins (1991:18-19) suggests that they can be used for standardisation (to make all students alike) or they can be used as tools for customisation (to help meet individual student's needs). There is value in setting meticulous high-level standards that teachers can expect all students to meet. As was indicated in section 2.5.2, *Criticism of traditional assessment*, the aim of assessment and standards should be to help individual students reach their maximum learning potential. High stakes testing (which is based on standardisation), as discussed in sections 2.3.2.1, *Standardised testing*, and 2.3.2.2, *Achievement testing*, does a great disservice to students. By aiming only at advanced level students and by concentrating on short answer, easy to mark questions, teachers forfeit the opportunity to maximise their students' learning potential.

It can, therefore, be concluded that more attention should be directed to understanding the principles that are behind good standards and quality assessment. Standards in education improve the effectiveness of achievement “by clearly defining what is to be taught and what kind of performance is expected” (Ravitch, 1995:25). Standards

should concern children's cognitive growth (Lewis, 1995:746). This important principle has been sorely absent from educational policies and practices in the past.

In order for standards to be implemented effectively, the principles for assessment must be closely related to the principles underlying standards. The principles of assessment have been discussed in section, 2.4, *Principles of assessment*.

There are several important principles that comprise the framework of a standards-based program.

### **3.4.1 Standards must be specific**

A standard has to be specific enough so that it can be measurable. If standards are too vague, students will not be able to know what is expected of them. Standards need to be clear, not only in their task description, but the task they are required to fulfill must lend itself to evaluation (Marzano in Scherer, 2001:17). For example, a benchmark such as “shows an understanding of World War I” doesn’t direct students to a specific, clear activity that demonstrates their mastery of this subject. A measurable task would be “explains the main reason why the United States entered World War I”. An additional benchmark might be “write a letter from a young soldier to his family conveying the events of a recent battle.” This benchmark requires students to use their knowledge of historical events as well as their letter writing skills.

Students need to know exactly what is expected of them so that they have the best chance of completing a task successfully. For example, a standard that asks students to understand a story is not specific enough. However, asking students to write a diary entry expressing how one of the characters felt at a particular point in the story would be a more specific and much clearer task description. Specific standards enable students to demonstrate their knowledge and skills and make it possible for teachers to assess their students’ work more accurately.

### **3.4.2 Standards should measure multiple levels**

Standards should reflect different levels of accomplishment, with many diverse levels

in each area of competence (Gandal, 1995:20). Students must have clear descriptions of the criteria they must meet on all levels and be able to understand how they can improve their level. They need to be provided with examples of signs or benchmarks that show that an objective has been accomplished, and there needs to be a division into proficiency goals at certain grade levels.

Standards have a direct impact on assessment. Standards must include multiple performance levels to allow for student diversity (Gandal, 1995:20). Gandal (1995:20) suggests that this can be done by developing multiple levels of achievement for each content standard, such as "proficient", "advanced" or "expert" levels that would permit students to reach their individual potential and monitor their progress.

### **3.4.3 Standards must be teachable**

For standards to really influence learning they must be teachable (Gandal & Vranek, 2001:7). This means that they must be clear and precise. Students need to know exactly what is expected of them in language that is easily understandable. Benchmarks, which are appropriate subcomponents, should describe the specific developmental components of the general domain identified by a standard. Gandal and Vranek (2001:7) point out that a standard that states that students should “read literally, inferentially, and critically”, fails to provide adequate direction to either teaching or learning.

Kendell and Marzano (1996:79-80) agree that benchmarks must clearly indicate standards that students can achieve. The following science benchmarks listed for K-2 (kindergarten through grade 2) in the USA within the standard, “Understands essential ideas about the composition and structure of the universe and earth’s place in it” could serve as examples:

A student should know:

- That the stars are innumerable, unevenly dispersed, and of unequal brightness
- That the sun can be seen only in the daytime, whereas the moon is out sometimes at night and sometimes during the day

- That the moon looks a little different every day, but looks the same again every four week

Wiggins (1991:19) agrees and stresses the important relationship between standards and assessment. He warns that standards become viable only when varied student performance and products meet them. He, therefore, calls for standards of “output” that would judge performance on authentic tasks.

Assessment needs to be able to measure the criteria set by the standards. Students need to be presented with performance opportunities that will enable them to use the knowledge and competence that they have acquired.

#### **3.4.4 Standards must be flexible**

Perhaps the most important principle is that standards must be flexible. Since the abilities of students vary, one can understand that there might be some standards that are unattainable for some students. In addition, since it is natural that students learn at a very different pace, it is understandable that there should be no rigid timetable for the attainment of all standards.

Covington (1996:24-25) explains that striving for excellence - that is, maximising the intellectual potential of each student - is the most legitimate of all academic goals and, happily, the one on which all interested parties can most likely agree. However, this kind of excellence is best advanced when achievement standards are implemented flexibly, according to the gifts and experiences of each child, not imposed uniformly for all children. If the goal of educational standards is to enable students to succeed, then students must be permitted to approach respective standards at their own pace while drawing on their individual skills and strengths.

Assessment should likewise be flexible to allow for greater student achievement. Flexible assessment leads to greater achievement by recognising the abilities and needs of a diverse student population. Students who have a choice of assessment

activities and who are allowed to work at their own speed will feel encouraged and motivated.

### 3.4.5 Number of standards

The number of standards must be reasonable. Marzano (in Scherer, 2001:15) points out that there is not enough school time to address all the standards; therefore the greatest obstacle to implementing standards is their current vast number.

The Educational Issues Department (EID) of the American Federation of Teachers (AFT) created their own list of 10 criteria which they consider to be the most important characteristics of good standards (Gandal, 1995:16-21). They stress that standards must be specific, manageable within the allotted time framework and apposite for assessing performance. The AFT maintains that too many standards cannot possibly be assessed. In addition, they recommend that standards include multiple performance levels, which combine various skills and knowledge. Although many of the principles behind standards are similar, there are different kinds of standards.

## 3.5 KINDS OF STANDARDS

One of the controversial issues surrounding standards is the question of what kind of standards should be applied. There are different kinds of standards, each having a different purpose and connection to instruction and assessment (Sandrock, 1997:8). The central question is whether standards should be **content** or **performance-based**.

**Content-based standards** describe standards in terms of knowledge and skills that should be acquired in order to exhibit competency (Lewis, 1995:746). Simply stated **content standards** reflect what students should know and be able to do. Assessment activities need to be designed with the content standards in mind. Students must be presented with challenging tasks that require them to utilise the skills and knowledge that they have acquired.

**Performance standards** define the level of learning that is considered to be satisfactory (Lewis, 1995:746). In other words, they describe how the students show they are accomplishing the **content standards**. Students are assessed through performance tasks in which they must apply and demonstrate what they know. They utilise their knowledge and skills in real life situations. This, we see is in sync with recent research findings (as discussed in sections 2.2.2, *Cognitivism*, 2.2.3, *Constructivism* and 2.2.4, *Multiple intelligences*). Some critics believe that **performance standards**, which are the most vital part of learning, have a close, inseparable connection to **content standards**. Marzano and Kendal (1996:14) point out that "performance standards identify the environments in which knowledge and skill should be demonstrated." They recommend that schools and districts begin by developing **content standards** and then define a "complementary set of **performance standards**" or performance tasks.

In addition, attention needs to be directed to **proficiency standards**, which supply the criteria and/or the scale for judging the degree of progress on the **performance standards**, and **program standards** that provide the descriptions of all the important components that must be covered when planning activities (Sandrock, 1997:8).

Although most discussions about standards generally center on content and performance standards, Lewis (1995:744-750) believes it is impossible to ignore the importance of **opportunity-to-learn standards**, and "**world class**" standards. **Opportunity-to-learn standards** relate to the conditions and resources necessary to help students meet performance standards. For example: high-quality teachers and sufficient instructional resources are needed in order to provide the circumstances that will enable students to exhibit their abilities and skills (Lewis, 1995:747).

"**World class**" standards are based on the **content standards** and expectations held for students in other countries, which have demonstrated superior performance on international assessments (Lewis, 1995:747). One of the rationales for establishing standards in America is the concern that students of other nations (Germany, France, Japan or England) are intellectually surpassing American students. It is, therefore, important to monitor standards around the world and emulate these high standards (Tucker & Coddling, 1998:45; Gandal, 1995:19).

### 3.6 PROBLEMS ASSOCIATED WITH STANDARDS

The initial controversy regarding standards is the question of whether standards should be implemented at all. Evers (2001:205-206) cites the major disagreement that continues between two major factions; the traditionalist school of thought and the progressive school of thought. The **traditionalists**, who believe in a methodical and orderly approach to learning that is teacher directed, favor the adoption of set standards. They support a type of learning that is based on a body of knowledge and set of skills, which they believe children should know. On the other hand, the **progressive** approach suggests that content should not be predetermined, but rather directed by a child's eager desire to learn (Evers, 2001:205-206). Yet, despite this controversy, the disillusionment with the existing educational system sets the standards movement in motion. Once standards-based education gained momentum it continued to be both praised and chastised.

Despite the general enthusiasm heralded by the adaptation of educational standards, experts remain skeptical. Sandrock (1997:7) warns that standards "hold tremendous promise, but also tremendous peril". There are innumerable reasons for questioning the effectiveness of standards. Firstly, Eisner (1993:22) advises that standards do not signify the most important ends we seek in education. He believes that we seek work "that displays ingenuity, complexity, and the student's personal signature". Cohen (1995:756) insists "Standards should be understood ... not as the kingpin of change or as the occasion to decide for our time what the content of education should be and what level of achievement will be acceptable." Rather they "should be understood as one tool for helping the entire education system to learn and improve..."

Maehr and Maehr (1996:23) warn that stressing standards is a quick, available solution to an issue that is considerably more complex. Though it is reassuring to accept that inadequacies in education can be easily resolved by applying standards, they are not the panacea they appear to be. According to Lewis (1995:745) standards will not save public education.

Although the standards movement holds potential, it has been met with a myriad of objections, which warrant attention in this paper.

### 3.6.1 Unclear standards

One of the most vehement criticisms directed at standards is the fact that they are often unclear (Schmoker & Marzano, 1999:21). As an advocate and engineer of national education standards in the Bush administration, Ravitch (1995:25-26) champions a clear definition of the skills that are to be taught and the specific performances standards that are aimed for. She calls for access to “equally challenging programs and course of study” for all students. Although she criticises some of the standards that have appeared, Ravitch acknowledges the need for standards. She views standards as an essential means towards ensuring that students in all schools have access to equally challenging education.

Even Ravitch (1999), who is a proponent of standardised-based education, bemoans standards that are vague and not precise. In this regard she cites examples, such as: “Students will read, write, listen and speak for critical analysis and evaluation” (a standard set for English language arts in New York City) and “Students will use a variety of intellectual skills to demonstrate their understanding of major ideas, eras, themes, developments and turning points in the history of the United States and New York” (a standard in social studies in New York City). These standards which she refers to as “higher, but hollow, academic standards” fail to provide any direction, and cannot help students achieve academic success (Ravitch, 1999).

Tucker and Coddling (1998) similarly criticise standards that are unclear. They point to the ambiguity of several content standards, such as “love of literature”, which offer little or no direction or indication of student ability. They suggest that performance standards are more suitable to measuring a student’s abilities in any given area. They view performance standards as a three-part system that encompasses:

- a) Concise descriptions of what students must know and be able to do
- b) Samples of student work to clearly demonstrate the standards that are required
- c) Commentaries on those samples that explain the characteristics that raise them to the standards

The implications of unclear standards for assessment are numerous. Without clear standards students will be unable to know what is expected of them. They may feel

frustrated and discouraged if they fail to meet the projected standards. Unclear standards may also result in poor assessment for students who have proficiency, but do not perform well as a result of poor direction.

### **3.6.2 Unrealistic standards**

Another criticism of standards is that it is unrealistic to assume that students who are already academically challenged can reach higher standards than in the past. Covington (1996:24) insists that if students cannot presently measure up to old, presumably less demanding standards, it seems pointless to make increased demands. Perhaps the primary issues related to national standards that have continued to elicit the most vehement criticism is the question as to whether it is feasible to expect all students to achieve the same proficiency.

It is essential to take into account that students' interests, motivation and opportunities to learn differ (O'Neil, 1993:7-8). As explained in section 2.6, *Alternative assessment*, students who are interested in a subject and can relate that subject to their own experiences will be motivated to apply themselves towards mastering the necessary skills (Eisner, 1991a:14). Therefore, standards must accept a range of performance activities and levels of performance to allow for students' differences.

An additional concern is that clearly defined standards might be too restrictive and unrealistic for learners who have limited abilities. Clearly not all students have the same aptitude. Moreover, not all standards offer students the opportunity to use *inter alia* their multiple intelligences to express what they know and are able to do. We have seen the importance of multiple intelligences in section 2.2.4, *Multiple intelligences*. It cannot be expected that all students will attain many of the standards if they are disallowed the chance to express themselves in ways in which they are most proficient. Another problem of applying uniformed standards is that since children come from diverse, socio-economic backgrounds some begin school at a disadvantage that widens over time. These children cannot make the learning associations that their classmates make. Furthermore, when standardised testing assesses them they are discriminated against even further (Popham, 2004a: 48). There are too many questions that appear on standardised achievement tests that children

from low-income families cannot answer. Popham (2004a:48) cites, as example, the question: “If you wanted to find out if a distant planet had mountains or rivers on it, which of the following tools should you use?”

- a) binoculars
- b) microscope
- c) telescope
- d) camera

The choices would be equally unintelligible for someone from a low-income group who had not been introduced to these tools in school. However, students from higher income families might watch science programs at home, discuss scientific topics with their families or have familiarity with some of these items (Popham, 2004a:48). In addition, new immigrants, as well as children with special needs such as learning disabilities, visual, auditory and attention deficit, may be penalised or not addressed correctly. Therefore, any initiative to determine standards should consider the needs and capabilities of all children in the education system.

Criticism is directed at the engineers of the standards movement for ignoring the contextual differences that children bring to school with them. It is admirable to want all students to reach the same goals, but the challenge remains as to how to bridge the gap suffered by students who have begun their education on a lower level than their peers (Popham, 2004a:48). Neill (2003:44) insists that standards alone cannot raise the level of education. He suggests that social investments and in and out of class support are needed to help students overcome academic gaps.

Unrealistic standards could impact negatively on assessment by providing an inaccurate appraisal of student achievement. Students who fail to meet these standards may be viewed as incompetent. Teachers who devote valuable time on standards that are impractical sacrifice attainable learning objectives. Their students may suffer if they are deprived of opportunities to face realistic goals.

### **3.6.3 Teaching to the test**

There is concern that the standards movement leads to test-driven instruction (Sizer,

1995:12). It is alarming to hear that one in three United States educators seems to be hyping test-focused teaching (Popham, 2004b:82). It is natural to assume that if teachers are made to feel accountable for test results they will feel compelled to dedicate more of their teaching time to the material dictated by the standards. As a result teaching will become less flexible and teachers will have less autonomy in their classrooms. The fact that standardised learning has connected itself to high-stake testing is a great problem. Too much emphasis has been placed on the importance of testing. Too much attention is being paid to the results of single grades which can restrict curriculum and limit instruction to just preparing for exams (Thompson, 2001:358).

Moreover, Wiggins (1991:23) warns that linking standards to test scores diminishes the value of standards by assuming that standards can be achieved through rote testing. Popham (2004b: 82-83) takes Wiggins' concerns even further. He objects to the practice of "item teaching" in which teachers drill on the specific material that will appear on the test. Rather than curriculum teaching which addresses the curricular aims directed by the standards, item teaching yields misleading results and, therefore, an incorrect evaluation of student performance. He cites as an example a list of two hundred words on a district approved spelling list. Teachers choose twenty words from the list, repetitively drill the students on those twenty words and then test them on the very same words. It would be incorrect to conclude that students who spelled the twenty words correctly would successfully spell all two hundred words on the list or that students who spelled only ten words correctly would successfully spell only one hundred words (Popham, 2004b:82). Tests are intended to demonstrate mastery of skills and knowledge represented by the test, and item testing does not do this (as explained in section 2.5.1, *Types of traditional assessment*).

- **Washback**

Alderson and Wall (1993:115) use the term *washback* (also called *backwash*) to refer to the idea that testing influences teaching. Biggs (1995:12) adds that *backwash* not only drives curriculum, but influences teaching methods and students' approach to

learning. Messick (1996:241) uses the term specifically to refer to how the use of tests influences language teachers and learners.

*Washback* has an effect on English teaching in Israel. Shohamy (1992:514) refers to *washback* as the use of external language tests to affect and drive foreign language learning in school.

*Washback* can have a direct effect on assessment. In Israel, when new EFL tests were introduced, teachers spent more class time with activities and tasks which were similar to those included on the tests (Shohamy, Donitsa-Schmits & Forman, 1996:306). Similarly, when new tests of Arabic second language (ASL) were originally implemented, teachers discontinued teaching new material in order to review and replaced textbooks with worksheets that mirrored the previous year's test (Shohamy et al., 1996:301). On the other hand, *washback* may ensure that minimal goals are met by all teachers in all schools so that children will be able to be accepted to any high school anywhere in the country.

Clearly, teaching to the test narrows the scope of learning and limits opportunities for authentic methods of assessment.

#### **3.6.4 Limiting expectations**

Other critics such as Olson (2000) blame the standards movement for narrowing the scope of knowledge acquisition, rather than the potential for learning. They criticise standards for encouraging limited expectations. Harnischfeger (1996:109), for example, points out that the exclusion of health, the arts and foreign languages in the list of goals in Bush's Goals 2000, manifests a lack of vision towards an educated America outfitted with the skills necessary to compete in a global society.

#### **3.6.5 Too many standards**

An additional, very powerful criticism is that the zeal of educators has resulted in an unrealistic abundance of standards. Finn (In Diegmueller, 1995:6) berates the policy makers and educators who produced standards drafts for having a penchant for

excessive and demanding aspirations. Brandt (1995:5) similarly complains that such ambitious standards are overwhelming. Ravitch (1995:5) compares the fourteen pounds and two thousand pages of standards documents in the United States to the three trim volumes of the Japanese national curriculum.

Marzano and Kendall (1998:1-5) fault the standards as being, not only too numerous, but repetitious and non-specific enough. Moreover, they accuse the massive number of standards of being the biggest obstacle to the implementation of standards. Their criticism refers to a study carried out at Mid-continent Research for Education and Learning (McRel) in Aurora, Colorado in the early 1990s whose objective was to build a plausible set of teaching and learning standards (REL Networks: 1999).

The research analysed one hundred and sixteen curriculum and standards documents at the state and national levels, which disclosed two hundred and fifty-six separate standards that relate to nearly four thousand accompanying benchmarks (specific topics) at various grade levels. They note that if American educators were to adequately cover all of the knowledge identified in the current set of standards for the core subject areas, it might take as much as twenty-two years of schooling as opposed to the current thirteen, within the current structure (Marzano & Kendall, 1998:3-5). Their suggestion is to decrease the number of standards, rather than increase the amount of instructional time. Clearly, an overabundance of standards can only be detrimental to the curriculum. From the preceding discussion it becomes clear that only a limited, agreed upon number of standards can have an effective influence on education.

Excessive standards have a serious consequence on assessment. They confuse learning goals and cloud the areas of proficiency that students should be expected to meet. When there are too many standards the most important ones may be sacrificed due to time restraints. Assessment, as a result, may focus on areas that are less significant for achievement.

### **3.6.6 Fails to consider the needs of diverse learners**

Educators like Kluth and Straut (2001:43-46) maintain that in order for standards to be effective and have significance, they must answer the needs of diverse learners. As discussed in section 2.2.4, *Multiple intelligences*, students can learn best when allowances are made for their different styles and abilities.

Kluth and Straut (2001:43-46) suggest conditions that will make it possible for all students to benefit from standards-based education. The conditions mentioned are:

- **Standards are developed and flexible**

This allows different students in the same classroom the opportunity to develop according to their own abilities, interests and needs. Whereas all students can work towards the same goals, the standards should allow a range of performance to recognise and accommodate student ability level. The following can serve as an example. Students understand the structure of a Shakespearean sonnet. They could be offered a variety of performance-based activities such as: a) explain the structure, b) compose an original sonnet or c) compare a sonnet that follows a different structure.

- **Standards require a wide range of assessment**

This advocates the implementation of data collection strategies such as portfolios, journals, and performances and presentations that would give a more meaningful analysis of students' ability and progress than traditional tests. These assessment tasks would be the most effective measurement of what a student knows and is able to do. Varying the type of performance or presentation can ensure that students have greater exposure to different techniques. For example, while a portfolio of student work could be required for the semester, a journal could be an end of the year requisite.

- **Standards allow equitable access to meaningful content**

It is believed that standards are the only present option to ensuring that students have

equal access to challenging educational objectives. Yet not all students are provided with stimulating curriculum that develops their capabilities. Teachers are responsible for exposing all their students to equal opportunities. Therefore, for example if computer access is unavailable for all students, resource material should be brought into the classroom for everyone's use.

### **3.6.7 Frustrating to the student**

An additional problem with the implementation of uniform standards is the very close association with high-stakes testing and accountability. Firstly, students are exposed to considerable pressure to succeed, which can cause students to feel frustrated and incompetent (Stiggins, 2001:12). Learning focuses on the learners' success on exams rather than on sustaining motivation and decreasing the number of drop-outs (Smith, 2004:84).

Furthermore, critics like Apple (in Diegmueller, 1995:56) believe that the standards movement discriminates against those who for various reasons generally perform poorly in schools. The quest to reach standards separates those who can achieve from those who cannot. Therefore, Apple (in Diegmueller, 1995:56) warns that national standards and national assessment are the beginning of attempts to segregate students.

The onus of standards can present other disadvantages for students. The race to meet standards may result in lengthening the school day, eliminating recess and demanding more homework, which can be an added cause of dissatisfaction and displeasure for students (Covington, 1996:24). Ratnesar (1999:56) agrees that teachers rely on homework so that their students can learn more. He echoes the fears that homework is a serious problem of increasing proportion.

The assessment implications are important to mention. Assessment methods should be customised to meet the needs of individual students, rather than standardised (Reigeluth, 1997:206). Students should be offered a variety of assessment choices so that they can have a fair chance to succeed. It is important to ease time restraints so that students who perform slowly can achieve the objectives that are placed before

them. Attention to these assessment principles will help motivate students and facilitate learning.

### **3.6.8 Drain on resources**

Popham (2004a: 48) suggests that there are schools that may be poorly equipped to offer equal learning opportunities. As a result, students in these schools will not be afforded the same learning opportunities as their peers.

Furthermore, the standards movement has been faulted for depleting resources that could be better used for more important needs. The expense invested on the standards movement could, for example, be directed towards implementing extracurricular activities, and more educational materials (Diegmüller, 1995:5).

### **3.6.9 Disservice to teachers**

The standards movement presents a disservice to teachers by failing to include them in decision making policies. In order for standards to be effective, teachers and educators need to be involved in the process. Unfortunately, standards are often developed outside the classroom in a “top-down procedure” (Smith, 2004:83). This shows disrespect and a lack of confidence in the very people who are entrusted with the education of our students. Moreover, by dictating what should be taught and assessed teachers are given the message that their discretion and judgment aren't valued (Hatch, 2002:459). Not only can this attitude lower teachers' self-esteem, it can also interfere with teachers' attitudes towards teaching and assessment. They may be inclined to invest less time and efforts in a profession in which they feel their talents and expertise are not acknowledged.

An additional problem is that teachers are ordered to teach and assess according to standards without receiving proper training and assistance. Since teachers are required to adapt a new approach for standards-based teaching; they need to adjust their teaching practices and change their methods of assessment. Lewis (1995:749-750) agrees that helping teachers implement standards requires extensive training, guidance and practice, which has not been sufficiently available.

Yet, despite the significant criticism of standards, not everyone has been dissatisfied with the actual standards that have begun to emerge. Standards have been received favorably by many educators. The following advantages of standards serve as examples to strengthen this viewpoint

### **3.7 ADVANTAGES OF STANDARDS**

The benefits of standards are mostly related to assessment practices. Therefore, there is a multitude of reasons to support standards-based education. Although considerable criticism has been directed against standards, the apparent benefits must be recognised.

#### **3.7.1 Equality of education**

The major advantage of standards-based education is rooted in the belief that if students are exposed to equal learning opportunities, they will have the same chance of success. We know that a large percentage of the student population begins their studies at a considerable disadvantage due to economic and social factors. In most, if not all cases, as students continue through the existing school system, the gap in their education widens. The idea of teaching and learning according to standards is intended to identify, isolate and then reinforce the standards that define educational achievements at various levels of learning (Ravitch, 1996). Therefore, students who are not up to specific standards can be more easily recognised and helped. Whereas there is the argument that standards discriminate against children who are poor and disadvantaged, Ravitch (1996) insists that standards are the vehicle, which will ensure that these children are taught the skills and knowledge that will help them secure higher education and better jobs.

Standards should take into account the diversity of learners and recognise that their abilities, interests and motivation are different (Smith, 2004:85). Yet, it is essential that students are assessed equally. This means that it is imperative that all students are assessed alongside the same standards of performance (Wiggins, 1991:21). In order to allow for student divergence, students should have numerous opportunities to

show that they have met the standards (Thompson, 2001:359). In addition, more time must be allotted for assessment.

### **3.7.2 Common standards and principles**

The incorporation of standards that are based on agreed upon principles is advantageous as it can supply common objectives and similar terminology for discussions about language instruction. Educators around the world, who share common standards, can share their teaching methods and experiences. Consequently, they can glean a wealth of knowledge from each other. Since teachers assume new responsibilities and challenges with the incorporation of standards-based education, the advantages of pooling knowledge and resources can enrich the teaching community tremendously (Resnick & Nolan, 1995:6-11). Schools in the process of developing standard-based curriculums can seek the guidance and support of international educational networks in choosing standards. Israel, for example, based its new curriculum in 2001 on principles and standards gleaned from international sources (Ministry of Education, Culture and Sport, 2001:7).

Standards have a direct influence on assessment as they are developed in order to inspire and evaluate performance (Gandal, 1995:20). Standards should describe what students are required to know and be able to do. Assessment activities should focus on creating opportunities for students to demonstrate their knowledge and skills. The most effective way to gather information on what a student knows and is able to do is to use a range of assessment activities, including portfolios, interviews, self-evaluation questionnaires, journals and learning logs (Lopez-Reyna, 1997:33-37).

Reigelruth (1997:204) advises that time restraints on the meeting of standards be dropped so that students can reach the same standards, but at different times. This means that assessment activities also allow students to complete projects or performances at different time.

### **3.7.3 A tool for better teaching**

There are several ways in which standards facilitate better teaching. Standard-based

teaching forces teachers to consider their learning objectives and to plan their lessons very clearly. They must have an accurate understanding of the principles they wish to transmit as well as a clear visualisation of the performance that is necessary to exhibit achievement of that standard. Therefore, assessment should form part of lesson planning right from the start.

- **Backward design**

According to the American Federation of Teachers (Jamentz, 2001:4-5), a key difference between standards-based and traditional instructional planning is the shift of emphasis from what students should know (**content standards**) to the quality of work every student will be expected to do (**performance standards**). Therefore, teachers must plan their lessons differently than in the past. The steps in lesson preparation which were previously suitable do not meet the needs of standards-based instruction. The teacher faces the challenge of planning backwards from an understanding of shared expectations for student performance to the lessons that will be required to assure that every student can achieve that level (Jamentz, 2001:4).

This shift has crucial significance for teacher preparation and professional development since teachers will need to approach lesson planning much differently than in the past. Teachers must have an unmistakable perception of the performances required to demonstrate achievement of the standards. They should develop meaningful assessment with their learning goals in mind prior to their lesson planning (McTighe & Thomas, 2003:52).

In the *backward design* process of Wiggins and McTighe (2000:8) teachers first consider the desired goals or standards they want to target. They then decide what evidence students should display in order to demonstrate that they have mastered the desired standards. Teachers next plan their teaching so that their lessons furnish their students with the necessary skills and knowledge that will enable them to perform successfully (Wiggins & McTighe, 2000:8).

The design process (Wiggins & McTighe, 2000:99) involves teachers in three stages of lesson planning, each with a central question.

- Stage 1-What is worthy of understanding?
- Stage 2-What is evidence of understanding?
- Stage 3-What learning experiences and teaching promote understanding, interest and excellence

According to their design, Wiggins and McTighe (2000:100) envision a learning process that extends beyond recall and encourage students to question, interpret and analyse what they learn. Teachers need to provide their students with lessons that expose them to inquiry, arguments, applications and areas that encourage deeper understanding and analysis (Wiggins & McTighe, 2000:99).

*Backward design* has a very significant effect on assessment. Since *backward design* encourages learning that extends beyond items that are easily measured by multiple-choice, possibilities exist for more effective and inventive assessment. Teachers have the freedom to develop a multitude of tasks that can appeal to varied interests and abilities. In order to assess students' understanding of the "full range of learning goals", it is important to have them apply their learning to new situations (McTighe & Thomas, 2003: 53).

In addition, *backward design*, clarifies performance tasks by supplying the student with scoring criteria and models of excellent performance as part of instruction (Wiggins, 1992:29). Scoring rubrics that are built with clear, descriptive language can guide students towards successful performance (Wiggins, 1992:30).

Assessment can be used to reflect on the success of standards-based instruction. Standards-based instruction can be clearly evaluated as it is based on evidence of satisfactory student learning as exhibited by work or performance that complies with or exceeds agreed-upon standards (Jamentz, 2001:13).

### **3.7.4 Raising and maintaining professional standards**

The United Federation of Teachers (UFT) embraces standards as a conduit for raising and maintaining professional standards (Gandal, 1995:16). Gandal maintains that when standards and monitoring practices are in effect, teachers and schools can be freed from conventionally onerous rules and given the flexibility to decide the best ways to help their students achieve at higher levels.

Sirotnik and Kimball (1999) and Reigeluth (1997:203) suggest that in order for standards to be effective, teachers must develop different assessment practices. This requires professional development that will be costly and ongoing. Sirotnik and Kimball (1999) go even further by stating that teachers need to be better compensated for the improvement in teaching and learning that they are asked to make.

### **3.7.5 A better learning experience**

According to Alexander (1993:10) a significant result of the standards movement is an invigorated classroom experience. The introduction of portfolios, projects and presentations as assessment methods has thrust classroom learning into a more creative, student-active mode, which can better develop and enhance learning.

### **3.7.6 The example of “world-class” standards**

Finally, as was explained in the introduction to this chapter (section 3.1), a strong reason for adopting a standards-based approach is that countries that have standards seem to be successful. Their success in academic arenas are worthy of emulation, and standards seem to be an objective that can be achieved (Gandal, 1995:19).

## **3.8 INFLUENCE OF STANDARDS ON ASSESSMENT**

Although the influence and implications for standards for assessment have been indicated throughout, standards also influenced assessment in the following ways. Standards require teachers to change their assessment in order to accommodate a diverse student population. They must offer an array of assessment activities to appeal

to different interests and abilities. Students do not learn at the same pace nor are they motivated by the same incentives. Assessment should present a wide range of opportunities for students to demonstrate what they know and are able to do with their skills and knowledge.

A major benefit of standards-based education is that national standards have resulted in the need for newer and improved assessment (Alexander, 1993:9). The same principles of more effective learning that are considered when creating standards should be used when developing assessment tasks. Standards should be created with assessment in mind. Gandal (1995:16-21) suggests a list of criteria for exemplary standards, which clearly illustrate the close connection between standards and assessment.

- Standards must focus on academics.
- Standards must be grounded in the core disciplines.
- Standards must be specific.
- Standards should be manageable within given restraints of time.
- Standards must be “world-class”.
- Standards must evaluate performance.
- Standards must include multiple performances.
- Standards must combine knowledge and skills.
- Standards must not dictate how materials are taught.

The symbiotic relationship between standards and assessment, as pointed out in section 2.8, *Relationship between standards and assessment*, benefits students; and helps guide teachers towards more effective and efficient language planning and instruction. The methods of assessment are as important as the standards as they both determine and direct learning goals. Van Rooyen and Prinsloo (2000:23) explain the relationship between assessment standards and assessment as follows: assessment standards set the guidelines for developing tasks to determine competence. Assessment standards are clear and transparent expressions of requirements against which successful (or unsuccessful) performance is assessed. They also capture the requirements for fair, valid and reliable assessment procedures.

As standards-based education became more popular, different countries around the world viewed standards as a way to solve their educational woes.

### **3.9 THE IMPLICATION OF THE MOVE TO STANDARDS-BASED EDUCATION IN A NUMBER OF COUNTRIES**

As previously stated standards-based education is not unique to the United States. Countries around the world regard standards as the answer to the educational problems that plague their schools. The creation of standards and the execution of programs to implement these standards have been tailored to the needs of educational communities around the world. In fact, the adoption of national standards in the United States was influenced by the academic success common to countries that have a national curriculum (as mentioned in section 3.1, *Introduction*). One of the reasons cited for the success of standards in education is the acknowledgement that students who understand what is expected of them will perform tasks and activities more efficiently and productively.

The numerous standards documents, which were published throughout the world, became models for countries that similarly searched for ways to improve their educational system. Standards in each country impacted on assessment practices.

#### **3.9.1 Standards in Queensland, Australia**

In 1989, Queensland, Australia published a national curriculum based on national standards. The standards have had a strong influence on assessment. Student assessment is promoted as an integral part of the teaching and learning process. Students are expected to assume responsibility for their own learning and self-monitoring (Sebba & Maxwell, 2005:192).

Assessment in Queensland, Australia is formative and relies heavily on higher order questioning (Sebba & Maxwell, 2005:200). The students and teachers develop shared learning objectives and students assess the learning that occurs in class. Considerable time is devoted to reflective analysis of the coursework. Feedback is important and

plays a pivotal role in assessment. Students assess themselves against the standards (Sebba & Maxwell, 2005:200).

### **3.9.2 Standards in Canada**

Curricular guidelines in Canada state learning objectives and standards that focus on the individual learning process (Sliwka, Fushell & Gauthier, 2005:97). The influence on assessment is apparent. Students help set their own learning goals and use portfolios and learning log books to monitor, record and reflect on their own learning (Sliwka, et al., 2005:97). Students and teachers confer on an individual, continuous basis. Self-assessment and peer assessment elements are integrated into the lessons so that students can develop critical thinking skills (Sliwka et. al., 2005:97). Assessment is criterion-referenced, rather than norm-referenced, which encourages the progress of each student (Sliwka et al., 2005:165).

### **3.9.3 Standards in England**

*The Education Reform Act (ERA)* of 1988 brought about wide-ranging changes in education by introducing a national curriculum and standards for compulsory schooling (Ruddock, 1995:81; Looney, 2005:129). The implementation of standards had a significant impact on assessment.

Assessment to support learning gained prominence in the early 1990s with the formation of the *Policy Task Group in Assessment* which supported and promoted the idea that formative assessment in classrooms resulted in significant advances in learning (Looney, 2005:130). The importance of assessment is apparent in schools in England. Lessons and assessment are transparent: teachers set up learning objectives at the commencement of class and students are provided with clear criteria for assignments, including homework (Looney, 2005:134).

Assessment directly involves students and has their interests in mind. Teachers share paradigms of work with their students so that students can recognise the level of achievement that they are expected to reach (Looney, 2005:134). In addition, students are given learning targets that are designed to meet their individual needs and

sometimes are required to assess their own work or the work of their peers (Looney, 2005:134-135).

#### **3.9.4 Standards in Finland**

*The Act for Comprehensive Education* (1998) outlines the national core curriculum for education in Finland and encourages the development of students' self-assessment skills (Voogt & Kasurinen, 2005:149). Assessment focuses on the development of the individual student, rather than on comparing students to each other. Feedback is considered essential for building self-knowledge and pupil motivation (Voogt, 2005:149). Assessment is seen as an instrument to sustain learning and to help pupils develop self-assessment skills that they will use later in life (Voogt, 2005:150).

#### **3.9.5 Standards in Italy**

The major reforms in the school system in Italy between the years 1997 and 2003 resulted in a move to develop new standards, tests and evaluation (Looney, Laneve & Moscato, 2005:163). Assessment is viewed as an instrument to improve learning, rather than to obtain a result (Looney et al., 2005:166).

Emphasis is placed on reinforcing formative assessment which has many benefits, such as helping to adapt learning to a diverse population of students (Looney et al., 2005:163). Time is invested in helping to develop self-assessment skills so that students can recognise progress and understand their mistakes (Looney et al., 2005:172). Daily assessment is used to support student progress and build self-confidence (Looney et al., 2005:173).

The standards movement, which has had a considerable influence on educational systems around the world, dramatically changed the approach to education and assessment in Israel.

### **3.10 STANDARDS IN ISRAEL**

Following the lead set by many countries around the world Israel has recently introduced standards-based education. Standards have been established in all major subject areas with the bulk of resources invested in the subjects of mathematics and English. Every child in Israel is required to study English in school as a foreign language (Hebrew is the official language of the country, which has a sizeable population of Arabs whose first language is Arabic). In addition, there are many immigrants in the country who speak English as their mother tongue as well as other languages, such as Russian, Amharit, Yiddish, French and Spanish.

English language instruction in state schools begins as early as the first grade (one to two hours a week) in some areas of the country although by grade four every child must attend English classes (three to five hours a week). The instruction of English is regarded as very important in the country. Graduate and postgraduate university programs in Israel require their students to be proficient in English. In addition, young people who complete their education and job training are expected to be able to communicate in English in all learning areas, including technology and science. This demands a high level of language proficiency; therefore, attention is directed towards improving English language study in Israel.

According to Smith (2004:83), however, the quality of language teaching in Israel has not brought students to the highest level of language proficiency attainable. As a result, considerable time and money has been invested in creating standards for learning English as a foreign language in schools. Similar to the situation in other places, the move towards standards has come about as a direct response to the criticism of the educational system in the country.

Whereas the teaching of English in the pre-1960 years was concerned with culture and literature, and the years between 1960 and 1990 focused on grammar and syntax, the 1990s shifted the stress on English as an international language of communication which would best prepare students for the 21<sup>st</sup> century. Motivated by an interest in elevating the level of language instruction in Israeli schools, educators and administrators have initiated major educational reform.

The first formal statement appertaining to the place of language learning in Israel appeared in 1995 in the document *Policy on Language Education in Israeli Schools* issued in the Ministry of Education Director-General's Circular dated June 1, 1995 (Ministry of Education, Culture and Sport, 1995) and, in a revised form, re-issued in April 1996 (Ministry of Education, Culture and Sport, 1996). The policy which took effect in September 1996, dealt with mother tongue teaching and second and foreign language education. The policy justified and revised existing policies and applications of language teaching. A distinct objective was to set general goals. In addition, the *Policy* recognised English as the country's first foreign language, thereby assuring the teaching of English as a foreign language in all Israeli schools.

### **3.10.1 The New Curriculum**

In July 1998, the Pedagogical Secretariat of the Ministry of Education, Culture and Sport approved a standards-based curriculum for the teaching of English in Israeli schools. The official draft copy of the new curriculum *Standards for Pupils of English: a Curriculum for Israeli Schools* was published in November 1998. This approach to include standards in the national curriculum was adopted after extensive research revealed that establishing and adhering to national standards in ESL teaching had been successful worldwide.

*Standards for Pupils of English* cited several documents and programs that had a significant influence on the new curriculum (Ministry of Education, 1998:7)

- The "CAN-DO" idea developed by John L.D. Clark (1978)
- The D.A. Wilkins' notional-functional syllabus (1976)
- The J.A. van Ek's Council of Europe Threshold program (1975)
- The T.J.M. van Els' Netherlands National Foreign Language Program (1992)
- The U.S. development of national standards for educational excellence as exemplified in the ESL Standards developed by the Center of Applied Linguistics for TESOL (Center for Applied Linguistics, 1997)
- The various versions of the Foreign Language Standards (National Standards in Foreign Language Education Project, 1996)

- The Department of Education and Employment (The National Curriculum for Modern Foreign Languages in the United Kingdom, 1996).

The new curriculum in Israel adopted standards that had been previously included in international standards documents. These standards included the following intended goals:

- 1) To provide equal opportunities for all pupils by judging them by identical standards, thus avoiding inconsistent teaching and assessment
- 2) To clarify what pupils needed to know at various levels of their education (Steiner, 2000:9)

Steiner (2000:12), the Chief Inspector for English Language Education, points to the similarities between the changes in the learning-teaching-assessment process in Israel and models around the world. She cites, for example, the influence of the Australian model where students assume an active role in knowledge acquisition (Steiner, 2000:12). The Australian Board of Education implemented the Victorian Certificate of Education (VCE), which is based on work requirements rather than testing. These work requirements, known as Common Assessment Tasks (CATs), comprise more than half of a pupil's grade and workload. They consist of writing reports, conducting research, making oral or audiovisual presentations, and additional activities that call for performance or production (Howe & Vickers, 1993:36-37). This influence has resulted in the heavy emphasis on task-based education and performance-based tasks in the new curriculum in Israel (Ministry of Education, 2001).

*The New Curriculum for the Teaching of English in Israel – Principles and Standards for Learning English as a Foreign Language in Israeli Schools* in 2001 was the first curriculum in English to be based on standards. This document will henceforth be referred to as *The New Curriculum*. Attention was given to defining the concepts and adopting standards that would meet national needs. In addition, arguments had to be presented clearly explaining the advantages as well as the disadvantages of a standard-based curriculum (Ministry of Education, 2001:9-11). Yet, the consensus of the members of the standards committee was that the benefits of such a significant undertaking would outweigh any negative aspects.

*The New Curriculum* (Ministry of Education, 2001:21) defines what students are expected to do at different levels of performance in four designated areas of language learning, which are referred to as domains of language ability and knowledge. The four areas are:

1. Social interaction
2. Access to information
3. Presentation
4. Appreciation of literature, culture and language

Each of these domains has clearly defined content and proficiency standards, levels of progression, benchmarks and criteria. The clearly designated benchmarks make it possible to identify levels of achievement. This classification differs from the previous syllabus that was based on the four skills of listening, speaking, reading and writing, which had dominated foreign language learning in the past in Israel. Replacing skills with domains allows for a more accurate representation of the goals and levels that have become the foundation of the curriculum for the teaching of English in Israel.

The committee that formulated *The New Curriculum* acted under the premise that setting standards would address the national need to elevate standards and to supply students completing an Israeli high school with the knowledge of English that is required by the modern world (Ministry of Education, Culture and Sport, 1998:10). Standards were chosen which would clearly designate the academic purposes of schooling. The **content standards**, it was decided, would need to apply equally to students of all races, ethnicities, and linguistic and cultural backgrounds. This clearly follows the precepts in international standards documents. The **proficiency standards**, it was decided, should define the level of proficiency that students demonstrate in the skills and knowledge framed by the state's content standards (Ministry of Education, Culture and Sport, 1998:10).

The committee strongly believed that the creation of standards would help clarify teaching goals. Moreover, they would help teachers focus on a specific body of skills and knowledge and on how most efficiently to utilise classroom time. Standards would also coordinate what is taught in different classrooms, building on skills and

knowledge from year to year. Yet, although there are clearly stated principles in the curriculum, teachers are encouraged to participate actively in developing material. They have the autonomy to choose material that is suitable for their particular student population (Steiner, 2000:9-12). *The New Curriculum* allows for teachers' creativity in designing assessment activities and in creating grading tools.

Furthermore, a set of principles of foreign language learning, education, assessment, and cognitive psychology and curriculum development were drafted. They were based on the results of studies on language learning (Ministry of Education, Culture and Sport, 2001:11). These principles are concerned with language learning, language teaching, choice of material, topics and tasks, and assessment. It was determined that language learning could be made more effective by providing students with maximum exposure to the target language and extensive opportunities for using the language.

Moreover, it is important for students to feel that they are making progress and see the usefulness of what they are learning. They need to be motivated to continue finding out about people, culture and literature connected through English. Language teaching, clearly, could be more efficacious in a language-rich environment in which there are opportunities for maximum verbal and visual stimulation. The goal should be to encourage the use of language outside the classroom and to broaden horizons through the use of English.

Teachers are encouraged to consider the pupil's level of cognitive and linguistic development and be aware of student diversity. The importance of these considerations was discussed in section 2.2.2, *Cognitivism*. Teachers should create a supportive classroom environment, where pupils are willing to seek knowledge on their own (Ministry of Education, Culture and Sport, 2001:15-18). The principles were arrived at after studying research carried out on foreign language theory. They are based on standards and guidelines for teaching English language arts and foreign language that have already been established abroad and proven successful in language instruction (Ministry of Education, Culture and Sport, 2001:10-11).

The publication of *The New Curriculum* became the official curriculum for all grades. This significant publication and its subsequent implementation have heralded an

exciting era in the area of second language teaching in the country. It holds tremendous promise for improving the achievement of language students. Yet, at the same time it presents a daunting challenge for teachers, who must meet the new demands of standards-based instruction and assessment.

### **3.10.2 The effect of standards on assessment**

The standards movement in Israel has had a significant impact on assessment. Steiner (2000:11) summarises the implications of a standards-based curriculum on assessment which include:

- Both formative and summative assessment
- A diversity of assessment methods
- Assessment tasks which permit pupils to display their knowledge, including the criteria for assessment
- Developmental assessment showing progress toward reaching the standards
- Assessment tasks which are included within the learning-teaching-assessment process and do not only comprise a final product

Whereas *The New Curriculum* was published to guide the curriculum, an additional document was published in order to help EFL teachers become better aware of the importance of standards and to raise the professional level of EFL teachers. *Professional Standards for English Teachers: Knowledge and Performance*, published in 1994 by the Ministry of Education Culture and Sport, set standards for the training and evaluation of English teachers in Israel. This document will be referred to in the rest of this thesis as *The Professional Standards*.

### **3.10.3 Teaching standards**

*The Professional Standards* (Ministry of Education, 2004:7) is organised into sections that outline the standards that have been set in each of the five domains as well as both the teacher's knowledge benchmarks and teaching performance benchmarks that must be met in order to meet each particular standard.

Domain	Standards
Content	<ul style="list-style-type: none"> <li>• Language Proficiency and Awareness</li> <li>• Literature and Culture</li> </ul>
Learning and the Learner	<ul style="list-style-type: none"> <li>• Theories and Practice of Learning and Learning an Additional Language</li> <li>• Individual Diversity</li> </ul>
Teaching and the Teacher	<ul style="list-style-type: none"> <li>• Classroom Interaction</li> <li>• Planning</li> <li>• Teaching Materials</li> <li>• The Teacher as a Professional</li> </ul>
Assessment	<ul style="list-style-type: none"> <li>• The Role of Assessment</li> <li>• The Methods of Assessment</li> <li>• The Learners' Role in Assessment</li> <li>• The Role of Testing in Assessment</li> </ul>
Classroom Environment	<ul style="list-style-type: none"> <li>• Classroom Climate and Management</li> <li>• The Physical Learning Environment</li> </ul>

The attention to standards in EFL teaching in Israel has had a considerable impact on assessment. Standards have not only affected the curriculum and the program of instruction, they have demanded a number of important changes in both teaching and assessment. The most dramatic changes are:

1. Assessment in the past was based on the behaviorist view of learning and tested pupils on determined material. The recent shift to a cognitive and constructivist approach calls for pupils to apply the knowledge and skills that they have acquired (Steiner, 2000:11-13).
2. As standards and subsequently assessment began to focus on the learner and his individual needs, assessment methods needed to adapt to this issue. Therefore, norm-referenced assessment, which measures a pupil's performance as compared to other pupils needed to be replaced. Criterion-referenced assessment, which measures a pupil's achievement according to a standard or standards, is more compatible with student-centered learning.

3. In a standards-based curriculum, the learner's progress and development are considered as essential elements of learning as well as the final product. For that reason, assessment needs to be continuous (formative) as well as summative.

*The Professional Standards* sets four clear standards in the domain of assessment (2004:16-19).

**Domain:** Assessment

**Standard:** The Role of Assessment

Teachers are aware of the role of assessment for improving learning as an integral part of the teaching-learning process. Teachers assess the progress of their learners as part of their teaching routine

**Teacher's Knowledge Benchmarks:** Teachers will meet this standard when they know:

- the interdependency of teaching, learning, and assessment
- the importance of providing feedback and monitoring as essential for effective learning
- the methods of assessment that take into account different levels, learning styles, and abilities in heterogeneous classes
- the value of alternative assessment

**Teaching Performance Benchmarks:** Teachers will meet this standard when they:

- integrate teaching, learning, and assessment in the planning of their units, lessons, and tasks
- ensure that learners are regularly aware of their language learning process
- design varied tasks and tools that allow learners to perform and succeed at different levels according to different learning styles and abilities
- use alternatives in assessment including projects, portfolios, presentations, exhibitions

**Standard: The Methods of Assessment**

Teachers know about theories and methods of assessment and match them with the appropriate tasks and tools.

**Teacher's Knowledge Benchmarks:** Teachers will meet this standard when they know about:

- the distinction between formative and summative assessment
- the various types of assessment methods such as performance-based tasks, projects, portfolios, tests, and the goals they are designed to achieve
- a wide range of assessment tools such as rubrics, assessment lists, and rating criteria to evaluate learners' achievement of the different goals that were set

**Teaching Performance Benchmarks:** Teachers will meet this standard when they:

- define for themselves and make clear to the learners the goals and criteria of the assessment task, prior to assigning it
- design different assessment tools that are developed with the participation of the learners and are according to the goals that were set
- provide a varied range of assessment tasks
- collect and record information about the learners' progress over a period of time from different sources including homework, assessment tasks, individual, pair and group activities

**Standard: The Learners' Role in Assessment**

Teachers are aware of the importance of involving learners and actively engaging them in the different stages of the assessment process.

**Teacher's Knowledge Benchmarks:** Teachers will meet this standard when they know:

- learning is enhanced when learners feel ownership of the assessment procedures
- about different assessment tools that learners can generate and use to assess their learning such as peer and self-assessment

- about assessment tools that allow learners to evaluate both process and product of their performance

**Teaching Performance Benchmarks:** Teachers will meet this standard when they:

- encourage learners to contribute to the design of the assessment procedures such as determining criteria, creating rubrics, and writing test items
- provide opportunities for learners to assess each other and themselves
- encourage the use of assessment tools for learners to evaluate process and product, such as checklists and rubrics

**Standard:** The Role of Testing in Assessment

Teachers know about the theories of language testing and design, and use tests appropriately:

**Teacher's Knowledge Benchmarks:** Teachers will meet this standard when they:

- know about criteria for the design of test and other assessment methods such as validity and reliability
- know about the practical constraints in designing and administering tests
- know about appropriate ways to test and assess the skills and domains according to the principles in the English Curriculum
- know about a wide range of types of test items such as multiple-choice, open-ended, true/false, their advantages and disadvantages, and when it is appropriate to use them
- are aware of test anxiety and its implications
- know about basic test calculations such as weighting, percentages, averages, distribution of grades
- know about national tests such as the *Bagrut* exams, and their implications for teaching

**Teaching Performance Benchmarks:** Teachers will meet this standard when they:

- design valid and reliable tests that are practical to administer and grade

- include test items appropriate to the goals of the test and the learning objectives
- design tests that have a balanced coverage of skills and domains as presented in the English Curriculum throughout the year
- take steps to lower test anxiety by appropriate preparation, and by creating a supportive classroom climate
- analyze test results using appropriate calculations
- take into account national tests in their teaching

The document outlines the core knowledge and performance that teachers should have mastered. The strong relationship between standards and assessment is very evident.

### **3.11 SUMMARY**

The intention of this chapter was to explore the powerful impact that the standards movement has had on education both internationally and in Israel. Though surely not a panacea, standards have produced the latest and most significant reforms in education. Although there have been considerable objections to adopting standards-based curricula, educators in countries around the world have eagerly embraced standards as a promising solution to the problems that haunt their educational systems.

Certainly the criticism directed at standards must be addressed. Firstly, the plethora of standards does present a significant impediment to the successful implementation of standards. Both teachers and students are daunted by an overabundance of demands. Clearly, too many standards cannot be taught nor reached.

Secondly, standards, which are inexplicit and unclear, are difficult or impossible to follow. Teachers cannot build their lessons around standards that are ambiguous and difficult to define. Students cannot strive to meet standards that they cannot understand nor follow guidelines that offer poor direction.

Thirdly, no student should be ignored in the quest to reach academic heights. Time and funding must be invested in order to fill the educational gaps that plague

disadvantaged students. In order for standards to be effective, they need to provide equal opportunities for all students.

Lastly, teachers need to understand the importance of standards; and the benefits that can be gained by implementing standards and standards-based assessment. Training is needed to equip teachers with the knowledge, skills and tools they require so they can teach according to the standards that serve as the foundation for language study and appreciation

The standards movement has had a considerable impact on assessment. Assessment activities need to be adjusted to reflect the principles of standards-based education, yet they must be adapted to meet the needs of a diverse student population.

Students are not the same. They come from different backgrounds, and they have different interests and abilities. Therefore, they should not be expected to all meet the standards at the same time. They should be afforded the opportunity to choose assessment activities that are suitable and work at a pace that is reasonable for them

In order for standards to be assessed, activities needed to be developed that can measure whether the standards were met. However, it is important to understand that not all students can reach standards at the same time. Assessment methods have to allow for flexible timetables that allow students to reach proficiency at their own pace.

As far as Israel is concerned, the movement to standards has implemented several changes in the approach to assessment. *The New Curriculum* has transformed English teaching in Israel. It has focused attention on assessment and placed the student at the center of the learning process.

The standards movement, which has had a profound effect on education and assessment, has revolutionised foreign language study. The next chapter will investigate the factors that influence the development of foreign language study and the direct effect on foreign language teaching and learning in Israel.

## CHAPTER 4

### FACTORS INFLUENCING THE TEACHING AND ASSESSMENT OF ENGLISH AS A FOREIGN LANGUAGE IN ISRAEL

#### 4.1 INTRODUCTION

Recent interest in cognitive theories has had an important impact on language learning. This influence is apparent in the new practices that have been implemented in second language (ESL) and foreign language (EFL) teaching. The most encompassing changes in recent years have been towards communicative language teaching (Nunan, 1999:9). In the 1960s language was viewed as a system of rules, and learners were expected to memorise and repeat these rules. In the 1970s, the move was to recognise language as a vehicle for expressing meaning and encouraging and maintaining communication. Therefore, the most effective approach to language should be teaching grammar and vocabulary communicatively, providing learners with experience in conversing and corresponding with real contexts in authentic settings (Nunan, 1999:78).

The changes in the purpose of language study have resulted in distinct changes in the role of the learner. In order to facilitate language proficiency, learners need to take charge of their learning. Therefore, education in general and language study and assessment specifically has become more learner-centered. As explained in section 2.5.2, *Criticism of traditional assessment*, learners who are more involved in the learning process are most willing to exert time, effort and resources in their own learning. It is also clear that motivation has a significant role in language acquisition (Ellis, 2000:200). A motivated student actively participates in learning tasks and stays on task without the need for constant direction or encouragement (Crookes & Schmidt, 1991:489). Attitude is also believed to have a direct influence on a student's motivation to learn language (Ellis, 2000:200). Students who see the direct benefits of learning language apply themselves more seriously.

According to Nunan (1999:14), the rate of language acquisition is increased considerably when learners have opportunities to interrelate with other language

users. Maximised involvement with other language users is beneficial as it considerably increases motivation, enhances performance, and cultivates interest in other cultures. Finally, researchers such as Chamot (2004:14-26) who have documented the findings of recent studies into the effect of language strategies on achieving learning goals, maintain that well developed learning strategies can foster language acquisition.

After years of admiring the success of foreign language programs in other countries, and wishing to emulate their accomplishments, the Ministry of Education in Israel has initiated major changes in the curriculum for English language study. This chapter intends to look into the history of language study and shed light on the theories and principles that promote and enrich language learning. In addition, this chapter reveals how cognitive influences have affected and shaped foreign language study around the world and in Israel.

As language teaching changed direction to assume a more communicative approach, this shift has had a direct effect on language assessment. Alternative ways of assessing learners were needed to monitor and check how pupils were able to utilise their language knowledge and abilities in authentic situations.

## **4.2 THE EFFECT OF GLOBALISATION ON ESL AND EFL LEARNING**

A global language is a language that is recognised for its particular communicative role in most countries around the world. This means that it generally is either an official language or the first foreign language in the language curriculum (Crystal, 1997:7-8; McKay, 2002:5). English is now the dominant or official language in over sixty countries and holds some position in every continent (Crystal, 1997:106).

There are many reasons for the global expansion of English. Political, military, and economic influences have been the primary factors (Crystal, 1997:7-8). In addition, the technological inventions and developments of the 21<sup>st</sup> century, such as telephone, fax, electronic mail, Internet etc. have facilitated worldwide communication in which the most frequent language used is English (Crystal, 1997:106). The fact that influential international organisations (such as the United Nations, the World Health

Organisation and the World Bank, for example) conduct business in English has contributed to making English a world language (Crystal, 1997:8-10).

English remains the international language of the world. Therefore, schools have the responsibility to empower students with the ability to communicate in English. In addition, since global education has become a growing worldwide trend it is now an important goal of education to help students become aware of issues that affect the world (Dyer & Bushell, 1996). The objectives of global education in language teaching are to enable students to attain and use foreign language in discussions about world issues and global problems (Cates, 1990:3).

Global education meshes very well with language teaching (Maley, 1992:73). Reading and writing about global issues bring real life considerations into the classroom and call for the use of authentic language for communication. Global education has an important effect on foreign and second language teaching. By involving students in discussions on world problems, teachers encourage students to use language for authentic purposes which promote language fluency.

In addition, global teaching is very much aligned with the principles of alternative assessment. Global teaching gives students a more active role in the classroom and creates a more student-centered learning environment (section 2.2.3, *Constructivism*). Global education challenges students to think about solutions to the world's problems. Critical thinking and finding solutions involves high level thinking that stimulates real understanding and learning (section 2.2.2, *Cognitivism*).

Whereas some countries feel threatened by the international dominance of English and view it as a danger to the local language, Prodromou (1999:4) suggests that English as a foreign or second language should function in conjunction with local languages and culture. This would help develop tolerance and provide students with a greater awareness of multiple cultures.

Globalisation has a strong influence on language study in Israel. English gained importance in the country for several reasons (Spolsky, 2002).

- Economic and political ties with the United States became more important.
- The growth of tourism to and from Israel made English a useful contact language.
- Radio, television, movies, books and magazines in English provided exposure to the global culture.
- There was an increase in immigration from English-speaking countries.
- English remains the main vehicle for access to the Web.

The challenge of globalisation for teachers of English in Israel is to prepare their students to be able to use English together with other languages as effective citizens in a multilingual world (Spolsky, 2002). Spolsky (2002) connects the issue of globalisation with the responsibilities and rights of English teachers in Israel. He believes that in order for teachers to function professionally and effectively, they should be provided with three essential things: a curriculum, preparation to teach that curriculum and a method of assessing the results. He recognises the development of the new curriculum, and the creation of guidelines for teacher training as positive steps in promoting language proficiency. However, he finds fault with the new Bagrut examinations (see section 4.15, *Matriculation Examination - Bagrut*). He argues that the modular system contradicts the idea that all features of language (speaking, reading, listening and writing) should be integrated.

Globalisation has affected English language study in Israel. Israel schools have participated in several co-existence projects, such as The Israel Middle East United Nations (TIMEMUN), The Garden of Peace: Association for the Communication of Bat Chen and The Israel UNICEF EFL Project. These projects encourage and promote global awareness and action in Israel (Lucas, 2005: 42-44).

#### **4.3 DISTINCTION BETWEEN THE TERMS ENGLISH AS A SECOND LANGUAGE (ESL) AND ENGLISH AS A FOREIGN LANGUAGE (EFL)**

The study of English for non-English speakers is generally referred to as either English as a second language (ESL) or English as a foreign language (EFL). According to Mitchell and Myles (1998:1) second language learning includes the learning of a language to any level under the condition that the second language

learning occurs sometime after the attainment of a primary language. Therefore, according to their definition second language is any language other than the learner's native language. In addition, they classify second languages as those languages, which one might be exposed to in various places in the community such as at work and/or the media. Crystal (1997:22) points out that there were 750 million EFL speakers in 1997 of which 375 million were first language speakers and 375 million were second language speakers.

Foreign languages, on the other hand, are languages that are not spoken in the local area and therefore, there is little or no exposure to those languages (Mitchell & Myles, 1998:1). These languages are usually studied in school for one or two hours a day in lower grades and three to five hours a week in higher grades.

ESL and EFL pupils need to be addressed differently since the classroom setting is very different for each group (Krieger, 2005:8-16). The pivotal factor that distinguishes each group is their accessibility to the target language (Brown, 2001:116), which means whether ample opportunities present themselves for language use outside the classroom. Brown (2001:75) connects language availability with the motivation to successfully use the language. Therefore, he points out that EFL students will have lower motivation to learn a language that is not spoken around them and which, they feel has little importance in their lives. They might, he suggests, have a desire to master the language in order to receive good grades or because succeeding in the language is a requirement for matriculation or further education.

ESL students, on the other hand, are subjected to more opportunities where they can use the language and understand its relevance to their lives. As a result they are motivated to apply themselves to learning the language and to becoming integrated into that language community (Irie, 2003:88). Therefore, EFL teachers need to create opportunities for students to use language in authentic settings. They need to motivate students by introducing them to activities where they can see the importance of language use (Brown, 2001:76).

English is taught as a foreign language in Israel. All schools in the Arab sector use Arabic as their language of instruction, and teach Hebrew as a second language and

English as a foreign language. In state schools, Hebrew is the language of instruction and English is taught as a foreign language (Spolsky & Shohamy, 1996). It is taught as a foreign language because English is not considered an official language in the country. There is, however, considerable direct exposure to English in Israel as a significant percentage of the population regularly speaks English. Signs in public offices, buildings, and on the roads as well as many official documents are printed in English. In addition, many areas of study in the universities, such as in the faculties of law, social work, psychology and medicine etc. are conducted in English. Israel is also permeated with innumerable television programs, radio broadcasts, newspapers, and cultural events in English.

ESL and EFL learners can both benefit from research that has been conducted on the subject of language acquisition. In this study the main focus is on EFL, but in certain sections reference is also made to ESL because they both relate to an assimilation of the English language.

#### **4.4 HISTORY OF SECOND AND FOREIGN LANGUAGE LEARNING**

It is important to review the history of second and foreign language study in order to understand how foreign language teaching and learning has developed.

##### **4.4.1 Introduction**

As discussed in section 2.2.1, *Behaviorism*, teaching in the 1950s and 1960s was in accordance with behaviorist theories, which stressed rigid adherence to rules and structure. Learning was based on the idea of stimulus and response, and the conviction that learning was the formation of habits (Mitchell & Myles, 1998:23).

There are apparent limits to this approach. In order to encourage an interest in language and stimulate enthusiasm, students should be exposed to authentic situations at an early stage so that they can see the benefit of using the target language. Clearly, in order to motivate learners to want to use the language they need to feel that they have the ability and opportunity to engage in real language situations. By delaying authentic communication until later stages of language learning as was done in the

past, students may become discouraged and dissuaded. Moreover, reluctant learners might be too daunted by what they perceive as the rigidity of syntax rules to express themselves in the new language. As a result, valuable opportunities for communication are lost.

#### **4.4.2 A change in focus**

In recent years, however, foreign language teaching has shifted its focus from teaching language as a set of skills (grammar, vocabulary and pronunciation), which could be evaluated by multiple-choice, true/false and short answer questions to teaching students to communicate in the target language in real-life situations (Thompson, 2000:3). Assessment in turn has changed from marking correct responses to linguistic questions to monitoring, measuring and assessing student performance in authentic language situations.

The shift in the 1970s towards a communicative view of language was a major step in promoting language use. Once the objective of language study focused on facilitating communication, the emphasis on adherence to rules of spelling, grammar and syntax needed to be replaced by skills more suited to verbal exchanges.

In order to encourage and develop language usage, emphasis should be placed on using language creatively and in authentic situations. This is because the focus on the creative character of language rather than on structural linguistics makes language accessible to more people, more alive and more valued (Mitchell & Myles, 1998:25).

Changes in the perception of how people learn language was not the single influence on foreign language study. The standards movement, which ushered in the dawn of modern education, has had an important effect on foreign language study as well. Switching to the implementation of language standards has resulted in the movement of language testing to assessment of actual language performances intended to help ESL and EFL students acquire a mastery of the English language.

#### **4.4.3 The effect of national standards on ESL and EFL teaching and assessment**

The implementation of national standards of education in the United States has had a dramatic impact on the teaching and assessment of English as a second or foreign language. Standards were created in language programs around the world to outline levels of proficiency that students were expected to achieve when learning foreign language. The move to language standards quickly gained acceptance.

A national survey on K-12 foreign language education sponsored by the Center for Applied Linguistics (CAL) in 1997 revealed that thirty-seven percent of the elementary school respondents with foreign language programs in their schools were aware of the *National Standards for Foreign Language Learning* (1996). Fifty-seven percent of the teachers said that their school's foreign language curriculum had changed because of an awareness of the standards. In secondary schools sixty-two percent indicated that they were aware of the *National Standards for Foreign Language Learning* and fifty-six percent acknowledged that the foreign language curriculum at their school had changed as a result of their awareness of the standards (Solomon, 1997). By the year 2000 more than forty states or districts had developed their own foreign language standards (Thompson, 2000:4).

#### **4.4.4 The American Council for the Teaching of Foreign Languages (ACTFL) Guidelines**

Many foreign language professionals believe that the *American Council for the Teaching of Foreign Languages Guidelines* (ACTFL Guidelines) created in 1983 by the American Council for the Teaching of Foreign Languages (ACTFL) was the impetus for the changes in the approach to foreign language instruction and assessment (Thompson, 2000:3). The United States government's language training department produced a set of guidelines that provided direction for assessing a student's language ability. These guidelines created specifically for second language learners recognise five distinct domains of performance: comprehension, language control, vocabulary usage, communicative strategies, and cultural awareness

The guidelines, which are specified for listening, speaking, reading and writing, identify stages of proficiency, rather than achievement. A student's language ability is measured according to four clearly defined levels of proficiency, from novice to superior, which are subdivided into nine subcategories that describe the specific skill of communication that the individual should be able to perform at each specific level (Shultz, 1998:8).

#### **4.4.5 The Standards for Foreign Language Learning – Preparing for the 21<sup>st</sup> Century**

The next significant advancement in foreign language teaching in the USA occurred in 1993 when a coalition of four national language organisations (the American Council on the Teaching of Foreign Languages, the American Association of Teachers of French, the American Association of Teachers of German and the American Association of Teachers of Spanish and Portuguese) was given funds to develop standards for foreign language education in grades K-12. In 1996 an eleven-member task force of foreign language professionals devised *Standards for Foreign Language Learning – Preparing for the 21<sup>st</sup> Century* which is a document intended to measure language proficiency. This document will henceforth be referred to as the *Standards*.

The premise behind the development of this document was the belief that “language and communication are at the heart of the human experience” (ACTFL, 1996). The underlying ambition of the *Standards* was to encourage all students to reach proficiency in English and at least one additional language (ACTFL, 1996). The intention of the *Standards* was, not to serve as a curriculum guide, but rather as a guide for suggesting the maximum and most effective ways of using authentic language. Whereas the *Standards* do not attempt to minimise the importance and essential role of grammar and vocabulary, they encourage teachers to direct their attention to the development of communicative ability as a whole, recognising the need for speakers of a foreign language to acquire the necessary skills that will enable them to communicate in a way that is suitable and meaningful with users of other languages (Thompson, 2000:3-4).

The *Standards*, which initiated wide-ranging changes in language teaching and assessment outlined the “five C’s”- communication, cultures, connections, comparisons, and communities - that are believed to be the framework for what all students of foreign language classes should be learning (Coltrane, 2002:6-8). The standards, which are identified by number, state what the language goals are in each particular area. These goals are identified by clearly defined objectives referred to as “progress indicators”, which are different and diverse performance-based activities. A student who reaches the respective standard in each goal area is expected to be able to master performance-based activities. For example: In the goal area of Communication (1.1) the following is a goal: “Students engage in conversation, provide and obtain information, express feelings and emotions, and exchange opinions.”

Examples of the progress indicators for this standard are:

- Grade 4: Students ask and answer questions about such things as family, school events, and celebrations in person or via letters, e-mail, or audio and videotapes.
- Grade 8: Students exchange information about personal events, memorable experiences, and other school subjects with peers and/or members of the target culture.
- Grade 12: Students exchange, support, and discuss their opinions and their individual perspectives with peers and/or speakers of the target language on a variety of topics dealing with contemporary and historical issues. (ACTFL, 1996).

The standards document further impacts assessment by providing examples of learning scenarios that are aimed at specific content standards (Leloup & Ponterio, 1998). The purposes of the learning scenarios are:

- To provide examples of the connection between classroom practice and standards
- To stimulate creativity in lesson and curriculum design
- To support student-based learning and standards-based instruction
- To show the interrelationship of the standards and their objectives

A significant development in foreign language instruction occurred in June 1997, when the teachers of English to Speakers of Other Languages (TESOL) released their own standards, *ESL Standards for Pre-K-12 Students*, in June of 1997. These standards define the content area for ESL by focusing on the language skills English students must master in order to function successfully in classrooms at all grade levels. In addition, it provides them with the language skills that they must master for productive management outside the classroom (Katz, 1999:1). These standards are worth discussing as they could also be applied to EFL.

There are nine ESL content standards that are arranged under three educational goals. They state what students should know and be able to do as a result of ESL instruction. An example of the ESL goals and standards are:

- Goal 1: To use English to communicate in social settings
  - Standard 1: Students will use English to participate in social settings.
  - Standard 2: Students will interact in, through, and with spoken and written English for personal expression and enjoyment.
  - Standard 3: Students will use learning strategies to extend their communicative competence.
  
- Goal 2: To use English to achieve academically in all content areas
  - Standard 1: Students will use English to interact in the classroom.
  - Standard 2: Students will use English to obtain, process, construct, and provide subject matter in spoken and written form.
  - Standard 3: Students will use appropriate learning strategies to construct and apply academic knowledge.
  
- Goal 3: To use English in socially and culturally appropriate ways
  - Standard 1: Students will choose a language variety, register, and genre according to audience, purpose, and setting.
  - Standard 2: Students will use non-verbal communication appropriate to audience, purpose, and setting.
  - Standard 3: Students will use appropriate learning strategies to extend their sociolinguistic and sociocultural competence (TESOL, 1997).

Each standard is explained with descriptors, sample progress indicators, and examples of specific classroom activities that support attainment of the standards. Although the standards represent student advancement towards attaining performance objectives, a wide range of standards represents the achievable level of each goal. This allows students to reach minimal goals and enables better students to surpass lower expectations. The advantage is that students who lack the abilities of their peers do have the ability to excel at their own level without feeling inferior to more adept classmates.

Assessment is developed to indicate the successful attainment of targeted standards. The wide range of standards allows students to reach different goals. In this way, assessment becomes more individualised and student centered. The clearly defined standards make it easier for students to understand the steps that they must take to attain the learning objectives.

These standards have the flexibility to allow teachers to consider them when preparing their lessons; yet do not bind teachers with any limits on their instruction or restraints. The flexibility of these standards avoids the criticism, which is often directed at standards, that standards restrict and limit teachers' freedom (section 3.6.9, *Disservice to teachers*). Whereas the *Standards* indicate what students should be taught in American foreign language classrooms, guidelines or "performance standards" were necessary to point to the proficiency that students could be expected to reach when attempting to master the content standards. These performance standards are outlined in the *ACTFL Performance Guidelines for K-12 Learners* (ACTF, 1998).

In order to direct attention to assessment of the new standards, ACTL published *ACTFL Performance Guidelines for K-12 Learners* in 1998. These guidelines center on second language use by K-12 students in foreign language programs. The guidelines recognise the continual cognitive development of language learners that affects their ability to execute language tasks (ACTFL, 1998). The assessment guidelines are intended to help teachers plan realistic assessment activities that are focused on the communicative function of language rather than on specific skills (ACTFL, 1998). They are divided into three modes of communication: Interpersonal

(active negotiation of meaning and conversation), Interpretive (cultural interpretation of meaning) and Presentational (the creation of messages when direct negotiation is unavailable). Performances are evaluated on the basis of comprehensibility, comprehension, language control, vocabulary, cultural awareness and communicative strategies.

#### **4.4.6 Managing the Assessment Process (The MAP)**

While the *ESL Standards* provided an entry point for including English language learners in the standards movement, it is just one of several TESOL efforts to reform ESL education. Another enterprising endeavor has been to sponsor the development of an assessment document, which serves as an accompaniment to the standards. This document, *Managing the Assessment Process: A Framework for Measuring Student Attainment of the ESL Standards (the MAP)* which was published in 1997 in the United States by TESOL had a significant influence on the assessment process (Katz, 1999:1). Instruction and assessment are entwined since not only do standards influence the choice and development of lesson plans and resource materials, they also have a direct connection to the pattern and choice of assessment methods. Therefore, *the MAP* intended to advance standards-based language proficiency by interweaving instruction methods and assessment practices (Katz, 1999:1).

The *MAP's* value is in its clarification of the distinctions that should be used in assessing student progress. Whereas a test is intended to measure skills or progress at a specific point in time, assessment is wider in range and entails collecting information over a span of time (Katz, 1999:2). An integral part of the *MAP* considers the planning of performance tasks and assessment aimed towards identifying the purpose and audience to which it is directed. It calls for the setting of performance levels (benchmarks) based on the learning goals. In addition, it encourages the judging of performances by using appropriate rubrics and scales and recommends evaluating progress over time with the aid of various and diversified assessments. According to Katz (1999:3) teachers need specific skills in order to conduct formal assessment that can meet minimum standards of reliability and validity (see sections 2.4.1, *Reliability*; 2.4.2, *Validity*).

The skills, which Katz recognises as imperative, include:

- Observing, interpreting and documenting learners' use of language
- Designing classroom tests and assessment tasks
- Providing diagnostic feedback to learners
- Evaluating the quality of test and assessment tasks
- Evaluating the quality of learners' language performance according to a rating scale
- Writing evaluative reports for program administrators

From the skills mentioned above, it can be concluded that assessment in language education needs to focus on language use. Assessment should directly involve the learner by providing a rating scale and feedback.

The policy documents mentioned above had an enormous influence on the development of foreign language standards in other countries. This influence is also apparent in the new curriculum in Israel, which outlines and explains the principles and standards for learning English as a foreign language for all grades in Israel.

#### **4.4.7 Standards and EFL in Israel**

Although many foreign language documents influenced the formation of *The New Curriculum*, the intention of the document is to address the particular needs of students in Israel (Ministry of Education, Culture and Sport, 2001:7). The purpose of a new curriculum is to set standards of excellence for the teaching of English as a foreign language in Israeli school by defining what is expected of students at different levels of performance in four areas (domains) of language learning (see section 3.10, *Standards in Israel*).

Section I of *The New Curriculum* (Ministry of Education, Culture and Sport, 2001:15-20) lists the principles for language study. These principles were gleaned from contemporary research in the fields of foreign language learning, education, assessment, cognitive psychology and curriculum development (Ministry of Education, Culture and Sport, 2001:12). The intention is that by applying these

principles to foreign language study students will learn language more effectively in a learning environment that encourages progress and attainment (Ministry of Education, Culture and Sport, 2001:12). The principles are intended to help students reach language proficiency. The following categories of principles are listed:

**a) Principles underlying language learning**

- Development of literary skills in the first language
- Maximum exposure to the target language and opportunities for using it
- Motivation and a willingness to invest effort and determination needed for foreign language learning
- Development of a positive self-image in the target language
- Development of confidence in the ability to use the target language
- Consciousness of how to learn and how to constantly develop better ways of learning the language
- Analysis and reflection on learning

**b) Principles underlying language teaching**

- Creating a language-rich environment
- Encouraging students to use English outside the classroom
- Stimulating students to broaden their horizons through the use of English
- Creating a supportive environment, one that allows students to take risks, to make errors and experiment with language

**c) Principles underlying the choice of materials:**

- Are appropriate to students' interests, experiences and knowledge
- Provide opportunities for meaningful communication
- Are presented in a variety of text types and media and are used for a variety of purposes
- Expand students' world knowledge by exposing them to relevant and current events

**d) Principles underlying the choice of content**

- Is unbiased, unprejudiced, inoffensive and non-stereotypical

- Appeals to a variety of backgrounds – religious, cultural and ethnic – and varying interests
- Stimulates students' interest in extensive reading, in the enjoyment of literature and in out-of-class value of English

**e) Principles underlying the choice of tasks**

- Are transparent to the student in terms of goals and process
- Encourage the use of English as a means for gaining information in other subjects
- Broaden students' horizons and motivate them to continue finding out about people, cultures, music and literature connected through English

**f) Principles underlying classroom assessment**

- Formative and summative assessment
- Assessment that allows for different levels of progress in language development
- Multiple methods of assessment for measuring language ability
- Feedback based on a collection of evidence from a variety of sources
- A wide range of assessment
- Criteria for assessment is known to pupils prior to the assessment
- Criteria for assessment can be negotiated between students and teachers
- Students take an active role in the assessment process
- Students are given ample time to revise work to be assessed
- Opportunities for peer and group assessment

From the preceding information it is clear that the principles underlying language teaching that have been incorporated into EFL study in Israel have clearly been influenced by research into EFL teaching and learning, standards-based instruction and newer, more innovative assessment methods. The fact that principles underlying assessment were included had certain implications. Teachers were instructed to adjust their teaching practices to incorporate principles of assessment that included formative and summative assessment using both traditional and alternative methods of

assessment. They were directed to focus on assessment that required students to demonstrate their skills in knowledge in meaningful situations.

The implications on learning are significant. Teachers are better able to formulate their academic goals when they have clear standards that serve as guidelines. Students have better opportunities to reach achievement objectives when their progress is monitored on an ongoing basis. They also have additional opportunities to improve work in progress.

However, progress in foreign language study has had other significant influences on language learning and teaching.

#### **4.5 FACTORS THAT INFLUENCE LANGUAGE LEARNING**

There are students who seem to learn language with relative ease. Therefore, attention has been focused on exploring characteristics and habits that distinguish these students from their less successful peers.

There has been significant discussion on the habits of language learners (Mitchell & Myles, 1998:18-19; Oxford, 1990a:10; Tremblay & Gardner, 1995:507). Just as certain students are likely to excel in certain subject areas, the question has arisen regarding whether there are certain types of students who have an inborn propensity for language or an ear for language. In addition, the ability to learn more than one graphic symbol for a sound and more than one set of grammatical rules also influences language learning. There are, however, several factors that seem to wield a considerable influence on language success.

##### **4.5.1 Affective factor**

Firstly, the affective factor according to social psychologists is the belief that the learner's attitude towards the second (L2) language and towards its speakers and culture can affect the learner's success or lack of success in mastering a language (Mitchell & Myles, 1998:18-19). Clearly, when students feel that they can garner benefits from the target language they will be more inclined to apply themselves

towards learning the language and communicating with native speakers. This interaction with native speakers breeds a genuine curiosity about their customs and culture. This curiosity can spark an interest in reading about different countries and engaging in additional activities related to the foreign language.

Assessment should, therefore, focus on authentic tasks where the learner clearly sees the benefits of acquiring the target language. These types of tasks should provide genuine communicative opportunities in which students can apply their language skills. Teachers need to create authentic activities that encourage communicative ability and that develop assessment methods that can accurately measure their students' communicative proficiency. These activities should focus on utilising language in situations that encourage language students to interact with native speakers.

#### **4.5.2 Student attitude**

Oxford (1990a:10) believes that student attitudes and behavior have an important influence on learning. She believes that students are basically interested in doing what is necessary in order to receive good grades, rather than in advancing valuable skills. Therefore, Oxford (1990a:10) advises teachers to encourage their students to become interested in language learning. They can do this by making learning and assessment more learner-centered, and by appealing to individual learning styles (see section 4.8, *Language learning: individual learning styles and assessment*):

Language experts agree that student attitude has a significant effect on foreign language learners (Tremblay & Gardner, 1995:507). Research shows that potential learners are most successful when they believe in their own ability to reach a high level of proficiency in the future. Teachers need to help students nurture a feeling of self-confidence and a willingness to engage in language activities. Guiding students towards reacting positively to groups that speak the target language, as well as showing a willingness to accept their assistance can expedite positive feelings. In addition, student motivation and language competency are both believed to increase in classrooms that foster an appreciation for English-speaking cultures (Tremblay & Gardner, 1995:506).

Attitude has a significant effect on assessment. It influences how learners evaluate their language ability. It is important for teachers to create a non-threatening learning environment that supports self-improvement. Constructive feedback should be given as it can encourage learning and give greater importance to assessment.

Assessment methods should include clear, explicit instructions in language that students can easily understand. This will give students direction in achieving learning objectives and as a result will help them develop confidence in their ability to succeed.

### **4.5.3 Motivational factor**

The motivational factor in language learning is very important as it drives the learner to achieve a goal, make an effort to attain the goal, and feel satisfaction with the task and its completion (Gardner & MacIntyre, 1993:2). Motivation can be both extrinsic when learners are encouraged by teachers, parents and peers or intrinsic when the drive to succeed emits from within the learners themselves.

The significant effect that motivation has on learning has been discussed in section (2.5.2, *Criticism of traditional assessment*). Dornyei (2001:116) believes that educators need to work to encourage motivation and insists that an integral element of teaching efficacy depends on the teachers' ability to motivate learners. He views teacher behavior as "a powerful motivational tool" (Dornyei, 2001:120), and suggests that in addition to the obvious incentives of reward and punishment, there is a litany of additional motivational techniques that teachers can use, such as: creating an amiable and encouraging classroom atmosphere, and developing good teacher-student relationships. Good and Brophy (1994:215) are in agreement with this view and stress that in order to develop motivation there must be a relaxed, non-threatening environment where students feel safe to take chances.

Assessment has a significant effect on learner motivation. Instructional feedback that provides learners with the independence to do their work increases motivation (Williams and Burden, 1997:135). Constructive feedback encourages learners to reflect on their work in order to revise and improve it. Alternate assessment, which

encourages students to choose tasks, can increase motivation and support language learning. Students who choose to undertake a certain task, and invest effort and determination tend to achieve higher levels of learning (Pintrich & Schunk, 1996:16).

#### **4.5.4 Level of self-confidence**

There is a strong relationship between having a positive self-image and performing well on learning tasks (Williams & Burden, 1997:72). Learners who have little confidence in their own abilities have difficulty overcoming the new challenges they face. This situation can be exacerbated in the language classroom since language study is foreign and students usually despair easily. Therefore, the language classroom has the potential to threaten and intimidate students by its very nature (Swain, 2000:112; Dornyei, 2001:130). Swain (2000:112) maintains that language students can learn best in an atmosphere that is reassuring rather than intimidating. He stresses the role of the language teacher in fostering a calm, assuring classroom climate that can encourage self-confidence.

Dornyei (2001:130) agrees on the importance of self-confidence. He suggests several approaches to sustain and enhance the learner's self-confidence, for example:

- Teachers can offer a personal word of encouragement.
- They can reduce classroom anxiety by making the learning context less stressful.
- Students can all be made to feel that they are contributing to the task thus enhancing interest.
- Teachers can help students acquire positive self-conceptions by giving repeated opportunities for success.

Tremblay and Gardner (1995:507) share the view that self-confidence helps learning and maintain that potential language learners are more successful when they believe in their own ability to excel in the future.

Assessment activities should be directed at building up confidence and creating situations where mistakes can be made and language can be used without embarrassment (Williams & Burden, 1997:73). The assessment activities should attach importance to the student's contribution to the task and have realistic objectives that can be reached. Academic objectives have to be clearly and accurately defined so that learners understand what they are expected to do. Criteria must be clear so that they can be assessed (Stiggins, 2005:27).

Students will become more confident in their abilities if they are given responsibility for their own learning. They can build their confidence in classroom situations where teachers act as advisors rather than authority figures (Young, 1991:432).

#### **4.5.5 Anxiety factor**

Anxiety exerts an additional influence on a student's desire to learn (Hawkins, 1999:128). Language anxiety refers to the uneasiness experienced when a situation demands the use of a second language with which the individual is not entirely proficient (Gardner & MacIntyre, 1993:5).

According to Gardner and MacIntyre (1993:5), language anxiety is a characteristic that relates to an individual's tendency to respond in a nervous manner when speaking in the second language. The anxious learner, who perceives the L2 as an uncomfortable experience, withdraws from voluntary participation, feels social pressure not to make mistakes and is reluctant to attempt uncertain linguistic forms (MacIntyre & Gardner, 1991:112). It is suggested that these students will have lower levels of verbal production, will have difficulty in basic learning and production, will be less likely to volunteer answers in class and will be reluctant to express personally relevant information in a second-language conversation (Gardner & MacIntyre, 1993:6).

The anxious learner will by nature be reluctant to speak in class or discourage interaction with speakers of the language. In many cases apprehension may be severe to the extent that students experience physiological manifestations, such as self-deprecation and rapid heart-beat (Gardner & MacIntyre, 1993:5). Therefore, it is

imperative for teachers to help instill students with self-assurance in their language ability.

MacIntyre, Noels and Clement (1997:269) explain that anxious learners may focus their attention on their perceived inadequacies and their potential for failure, rather than on concentrating on the task itself. As a result, their ability to reach language proficiency will be impaired. It is important to use assessment methods, which arouse interest and minimise anxiety. Assessment that concentrates on communicative fluency rather than on linguistic accuracy can help make students feel more comfortable and more able to succeed (Shaaban, 2001:16). Assessment should emphasise learning as a continuous, remedial process whereby learners feel confident that they can improve their skills and increase their knowledge of the language.

#### **4.5.6 Feelings of autonomy**

Contemporary views point to the importance of focusing on the learner (Williams & Burden, 1997:205). This means encouraging learners to participate in course design, decision-making and assessment. The most effective learners set goals for themselves and self-assess their work (McTighe & O'Connor, 2005:11).

We have seen in section 2.5.2, *Criticism of traditional assessment*, that children become more avid learners when they are involved in planning and choosing learning tasks. Their involvement can vary from attention to small detail to a very active role in decision-making. The amount of responsibility can be increased over time, which in itself can serve as a motivational factor.

Boomer (1992:10-11) agrees that learners benefit by active involvement in their own learning. He suggests the following stages that learners can be involved in:

- Joint planning and negotiation
- Setting the aims (including teacher's and learner's intentions)
- Collaborative exploration
- Achievements

- Evaluation (shared reflection)

Even the most elementary instructions lend themselves to allowing and encouraging learner input. Learners can participate in additional decisions, such as: whether they favor individual, pair or group work, which due dates seem reasonable for assignments and the number of times that process writing assignments should be submitted for grading.

Nunan (1999:164-165) suggests ways to include students in decision-making. For instance, if a lesson objective is to utilise the information in a given chart to convey information, the students could be encouraged to choose the activity they wish to do. For example, Use the chart in Task 2 for ideas and complete one of the following activities: a) Talk about city life and country life in your country or b) Write your own article about city and country living. By allowing learners to choose activities, we are encouraging them to make linguistic choices (Nunan, 1999:165). In addition, they are given the opportunity to draw upon the intelligence that they feel comfortable and confident using in a particular situation. The importance of recognising the role of multiple intelligences has been discussed in section 2.2.4, *Multiple intelligences*.

Autonomous learning directly effects assessment. Autonomous learners not only select their tasks and develop them, they self-monitor their performance, self-evaluate and self-reinforce. The evaluation process involves three steps: 1) learners examine the result of their efforts to learn, 2) they assess the criteria that they will use to judge it; and 3) they apply the criteria (Wenden, 1998:28). Following these steps aligns the principles of assessment and learning which were discussed previously in this thesis (section, 2.6.1.7, *Improvement of learning*).

#### **4.6 THE INFLUENCE OF COGNITIVE THEORIES ON ESL AND EFL LEARNING**

This thesis (section 2.2.2, *Cognitivism*) has discussed the theories of cognitive psychology which conclude that successful learning requires learners to actively integrate new information into their existing mental structures, thereby developing their own understanding. In language learning, this enables foreign language students

to enrich their awareness of the second language and its contiguous culture (Oxford & McGroarty, 1990:57-58). The most significant implication of the cognitive connection to language learning is that by relying on cognitive principles, all learners (not only the most capable learners) can be taught strategies that can facilitate language acquisition.

O'Malley & Chamot, (1990:217) believe that there are several benefits that can be gained by applying cognitive theories to L2 acquisition. Firstly, learning can be an active process in which individuals utilise information and strategies. Secondly, the complex skills that are employed to store and learn other disciplines can be used in language learning. Thirdly, language awareness can accompany the learner through the various stages of language acquisition such as initial awareness, active involvement in using information and culminate with the satisfaction of full independent language use. Fourthly, the use of learning strategies and cognitive processes can result in successful learning outcomes.

Van Lier (1996:11) points out that "in order to learn something new, one must first notice it". Therefore, teachers are responsible for making students aware of the new language and help them to focus and process the language. According to Little (1991:45) learning is possible only to the extent to which learners are able to integrate new information in relationship to their experiences. In language learning, this integration can occur through exposure to the language and language engagement.

#### **4.7 THE GOOD LANGUAGE LEARNER**

There has been a growing interest in the "good language learner" - the learner who seems to master language with relative ease and comfort. A successful language learner makes effectual choices when choosing learning tasks and strategies (Nunan, 1999:193).

The previous assumption that successful language learners possessed an inherent ability to learn languages or had a special ear for language has been supplanted by a new understanding. The conjecture today is that language learners employ special resources that enhance learning. Furthermore, they assume a more active role in their

language studies. As a result, attention has been directed at isolating and identifying something special or unique that can be learned from successful language learners (O'Malley & Chamot, 1990:2). For example, they not only use strategies more frequently, they use a greater variety of strategies (O'Malley & Chamot, 1990:2). In addition, Vann and Abraham (1990:192) believe that they bring the strategies to completion and utilise them more correctly than poor language learners. This would explain why poor language learners who might also use good language learner strategies fail to master a language.

Teachers should use the information that they gather from identifying the attributes of good language learners to help other students who have difficulty in mastering language. In addition, they should introduce good language learning strategies into their language classrooms in order to maximise the possibilities for language achievement.

#### **4.8 LANGUAGE LEARNING: INDIVIDUAL LEARNING STYLES AND ASSESSMENT**

Individual learning styles can have an effect on language learning. If learning styles can be identified and adapted, students can learn language more efficiently and effectively. Furthermore, teachers can introduce students to learning styles that they may be unfamiliar with.

In order to ascertain the learning styles of individual students Lessard-Clouston (1997) and Sharkey (1994-5:19) suggest that teachers monitor student behavior in class, and provide them with short questionnaires inquiring about their language objectives and which classroom activities they particularly enjoy or dislike (Lessard-Clouston, 1997; Sharkey, 1994-5:19). This can give teachers considerable information about a student's goals and motivation. This knowledge can help teachers supply a broad range of strategies intended for diverse learning styles (Lessard-Clouston, 1997).

Nunan (1996:36) and Graham (1997:170) suggest that teachers encourage learner reflection by asking students to keep journals about what they learned, where and

when they used English, which errors they made and what they perceive to be their own language difficulties. This information can provide invaluable information that can help teachers direct students towards improving their language learning.

Different types of assessment are needed to cater to different learning styles. Teachers should allow for a varied array of assessment tasks that can appeal to different levels of student interest and ability. Individualised assessment is a valuable way to recognise student differences and encourage student achievement. At the same time it recognises the importance of the principles of multiple intelligences (section 2.2.4, *Multiple intelligences*).

#### **4.9 LANGUAGE LEARNING: STRATEGIES IN ESL AND EFL**

The influence of cognitive learning theories has resulted in greater attention to strategies in language learning. O'Malley and Chamot (1990:1) define language learning strategies as “the special thought or behaviors that individuals use to help them comprehend, learn or retain new information”.

Research into language learning strategies (LLS) can be traced to the 1960s, and find its base in cognitive psychology (Williams & Burden, 1997:149). These strategies are premised on the understanding that learners are not passive in their learning, but are actively involved in making sense of tasks with which they are faced in order to learn (Williams & Burden, 1997:144).

Current research clearly shows that the effectual use of strategies is linked to high levels of attainment in a second language (Gardner & MacIntyre, 1992:218; Graham, 1997:42-43). They improve language performance, encourage learner autonomy, are teachable and expand the role of teacher and student (Oxford & McGroarty, 1990:59). Therefore, significant attention has been directed in recent years towards identifying learning strategies that could facilitate language learning.

According to Lessard-Clouston (1997), there are basic features that define language learning strategies, such as:

- They are learner generated, they are steps taken directly by the learner.
- They enrich language learning and assist in the development of language proficiency as reflected in the learner's skills in reading, writing or listening.
- They may be apparent, such as behavioral skills or unseen such as thoughts and mental processes.
- They involve information and memory such as knowledge, grammar rules, etc.

#### **4.9.1 Advantages of LLS in EFL**

There are numerous reasons to support the development of language learning strategies.

- LLS exert a considerable influence on the degree of student motivation to learn a foreign language by facilitating language learning by empowering students with control over their own language activities. The best strategies are “tools for active, self-directed movement for developing communicative competence” (Oxford, 1990a:1).
- LLS involve the conscious plans used by the student thereby calling upon cognitive components that influence learning (Gardner & MacIntyre, 1992:219). Stern (1992:261) agrees that students themselves must initiate learning strategies. She believes that the very idea of learning strategy depends on a conscious attempt on the part of the learner to participate in activities directed towards accomplishing specific objectives. Therefore, any deliberate guidelines and learning methods would be classified as learning strategies.
- Strategy training is most effective when it is included as a natural component of the students' classroom experience (Oxford, 1992:19). Therefore, it is the responsibility of teachers to train their students in strategies that will advance language learning and to use strategies regularly in the classroom. Lessard-Clouston (1997:3) develops this idea even further when he states that assisting students to understand good language learning strategies is a valued

characteristic of a good language teacher.

- LLS help learners develop communicative competence (Oxford, 1990a:9). Since the objective of language instruction is to build communicative competency students should be encouraged to adopt strategies that will facilitate communication in the target language.

We can understand from the list above that LLS have implications for both teachers and students. The teacher's responsibility is to encourage students and prepare them to take charge of their own learning. Students need to take initiatives and assume an active role in language learning. The dynamic effect of students' involvement in learning has been documented earlier in this chapter (section, 4.5.6, *Feelings of autonomy*).

#### 4.9.2 Types of language learning strategies

There are different types of language learning strategies. Oxford (1990a: 17) divides them into two distinct groups; those that are direct and those that are indirect. The three types of **direct strategies** are: 1) memory, 2) cognitive strategies and 3) compensation strategies. **Indirect strategies** are: 1) metacognitive strategies, 2) affective strategies and 3) social strategies.

The focus on the individual learner is a reason to incorporate strategies based learning into the language classroom. As teachers in all areas of education have begun to step back to allow learners to assume control of their own learning, strategies can help learners understand and direct their learning.

It is the responsibility of teachers to introduce their students to those strategies that will best equip them to grapple with the demands of language learning. Teachers need to be interested in teaching language learning strategies. They need to understand their importance and be willing to invest time effort and resources in their development.

Teachers can utilise the information that is known about language learning strategies to help their students develop individual approaches to language study. These strategies can facilitate language proficiency and help students feel more confident learning language.

#### **4.10 THE ROLE OF FIRST LANGUAGE ON FOREIGN LANGUAGE STUDY**

A first language (L1) has a significant influence on the learning of a second (L2) or additional.

##### **4.10.1 Reasons for learning language**

Learning language is valuable as it encourages students “to step outside their mother tongue and see it in some kind of perspective” (Hawkins, 1999:128). Learners gain a better understanding of their native language when they are engaged in second or foreign language instruction. As a result, their knowledge of the first language and their proficiency in that language is further developed through the study of an additional language.

During the past two decades, foreign language teaching has shifted from a traditional grammar-based approach to more communicative and interactive approaches. This change has resulted in changes in the way language is taught (Bell, 2005: 261).

##### **4.10.2 Connection between first language and second or foreign language learning**

Although first language acquisition is a different activity from foreign language learning in school, there is a strong connection between first language and second or foreign language acquisition (Hawkins, 1999:133). Researchers like Padilla and Sung (1990:50) indicate that when two languages share many linguistic features (phonetic systems, syntactic structure and semantic usage), for example, Spanish and Portuguese or Korean and Japanese, there is a faster rate of language learning than when two languages are very different, for example English and Korean. The reason is that processing time is slower when two languages have dissimilar linguistic

features because learners must learn most of the linguistic features of a new language. On the other hand, similar languages might cause interference between the two languages that could also be a detriment by causing confusion (Padilla & Sung, 1990:53).

#### **4.10.3 Negative effects of first language on second or foreign language learning**

There has been a continuing debate over the use of the mother tongue (L1) in the foreign language classroom (Tang, 2002:36). It is believed that too much dependence on the mother tongue even when it is possible to try using the foreign language will impede achievement and competence (Oxford & McGroarty, 1990:73). The negative attitude toward L1 has resulted in the suggestion that mother tongue in EFL is “a skeleton in the cupboard” (Prodromou, 2000) and warns of the detrimental consequences of using LI in EFL instruction. Teachers who may lapse into L1 often feel a sense of guilt or failure regarding their language instruction (Auerbach, 1993:13).

Relying entirely on L2 can lead to alienation from the learning process (Puchler & Field, 2001:85). Students may become frustrated when they are unable to express themselves in the second language. They may feel disconnected from the lessons when they have information or experiences that they wish to share, but are impeded from participating because they cannot express themselves in L2. This frustration creates a barrier between students and teachers, which can be reduced or alleviated when L1 is permitted for clarification (Puchler & Field, 2001:86; Harbord, 1992:354).

Second language learners can experience problems when they have trouble understanding instructions in the foreign language. They may perform badly and fail to demonstrate their true skills and abilities. As a consequence, assessment will be clouded by language difficulties. There are additional assessment problems that second language learners encounter during their language studies. The language achievements of L2 students are usually measured against those of L1 native speakers (Pillar, 2002:180). Yet, such a comparison is not reliable. The competence and performance of L1 and L2 users cannot be measured according to identical criteria

since the L2 learner is multilingual whereas L1 users possess monolingual competency (Pillar, 2002:180).

However, Burden (2000:5) believes that L2 use should be maximised as the classroom often provides the students' only exposure to English. In addition, reliance on the use of L2 helps to demonstrate the importance of the second language (Puchler & Field, 2001:86).

#### **4.10.4 Positive effects of first language on second or foreign language learning**

Despite some apprehension, in recent years, several language experts have disclaimed the argument that L1 has a negative influence in the EFL classroom (Medgyes, 1994:66). On the contrary, there has been more attention directed to the positive role of L1 in the EFL classroom.

##### **4.10.4.1 L1 provides students with a sense of security and identity**

Students who study a foreign language often experience feelings of insecurity. Consequently, they may be reluctant to express themselves in the new language. Starting with the L1 affords a sense of security and encourages students to express themselves (Auerbach, 1993:19). Schweers (1999:9) supports the limited use of L1 if it makes students feel more secure. Using L1 helps ease the discomfort students might feel and diminish the negative feelings that students might have towards being required to learn the language (Schweers, 1999:9).

The feeling of security affects assessment since it is apparent that students who feel confident and comfortable in the classroom learn more and perform better (section 4.5.4, *Level of self-confidence*). Tang (2002:41) agrees with Schweers and suggests that using L1 in the English classroom can assist in the teaching and learning process rather than impede it. Ur (2003:64), a prominent EFL educator in Israel, supports limited and controlled use of L1 in order to provide clarity and confidence.

Clarity, as well as confidence plays a pivotal role in learning and assessment. Clearly, students must have clear instructions in order to perform to the best of their ability

(section 2.6.1.1 *Clear instructions*). When students do not understand instructions in L2, they are unable to carry out their assignments correctly and thus teachers cannot assess them accurately.

#### **4.10.4.2 L1 Supplies a foundation for language understanding**

It is important for both teachers and students to realise the importance of the influence of first language for second and foreign language learning. Students' expectations of the foreign language are formed by their experiences with the first language (Hawkins, 1999:133). In Hebrew, for example, the perfect tense does not exist. Therefore, understanding a past tense form that relates to the present tense is a very difficult concept for Hebrew speakers. When teachers establish the connection between the ways present tense is used in Hebrew (i.e. there is a simple present tense that is similar to the present tense in English, and the word "k'var" – meaning "already" is added before present tense verbs that show the connection of a past action to the present) for instance, they are better able to explain the role of the perfect tense in English. Explaining foreign language in relationship to the first language (L1) expedites understanding and makes acceptance more possible.

Krieger (2005:14) points out that students will inevitably use L1 in the classroom whether or not it is permitted. Therefore, he advises teachers to direct the use of L1 towards a pedagogically beneficial objective.

### **4.11 TEACHING AND ASSESSMENT IN THE EFL CLASSROOM**

New methods have been incorporated into EFL teaching in order to meet the needs of language students. Theories of how learners acquire language as well as an emphasis on the communicative role of language are the pivotal factors that are considered in language teaching.

The connection between instruction and assessment has been documented in this thesis (section, 2.5.2 *Criticism of traditional assessment*). Alternative assessment has become popular for EFL teaching since it supports the objectives of language acquisition. It affords students the opportunity to demonstrate what they are able to

do with language. The chance to use authentic language in real-life settings rather than in contrived classroom situations makes language learning more meaningful. Learners can benefit by using language and hearing language in a variety of contexts that simply don't naturally exist in a classroom setting. By requiring learners to engage in performance-based tasks that mirror real-life situations, teachers are allowing learners to produce authentic language.

There are several teaching methods that can facilitate language learning. These methods are based on several principles that should be considered when selecting EFL classroom activities (Krieger, 2005:10-11).

These principles are:

- Teachers much deal with the fact that there is little or no exposure to English outside the classroom (It is however, becoming more and more the case that the use of the Internet is providing wider exposure to English).
- Teachers need to maximise fluency practice in order to develop fluency.
- Teachers need to be judicious in the selection of activities to ensure the use of English.

Each of these principles is seeded in the belief that English needs to be used actively in EFL learning and assessment.

There are criteria for selecting EFL classroom activities (Krieger, 2005:11). In order to best educate EFL from students in the EFL classroom, an activity should:

- Have a evident, clear and convincing objective
- Have English use built into the logic of the activity
- Not be too cognitively demanding to manage English
- Be interesting to the student

The elements mentioned above reflect an approach to EFL learning that can best maximise language proficiency.

#### 4.12 ALTERNATIVE ASSESSMENT IN EFL

The important connection between assessment and learning (Gipps, 1994:18; Williams & Burden, 1997:42) has been documented in this thesis (section 2.5.2, *Criticism of traditional assessment*). EFL instruction and assessment should focus on the principles of assessment (section, 2.4, *Principles of assessment*) that advance language learning and proficiency. Assessment of language should examine what learners can do with language in authentic tasks and challenge students to engage in higher order thinking tasks.

It is apparent from this dissertation that a central aim of learning is to enable the learner to be autonomous (section 2.2.3, *Constructivism*). EFL teaching and assessment should encourage students to learn independently using their skills and knowledge. When pupils feel a connection to the learning material, they become more involved in what they learn (section 2.2.3, *Constructivism*). The EFL classroom should include authentic, relevant material that can arouse and sustain student interest.

Authentic assessment can improve language learning in many ways. Kohonen (1999: 292) suggests that alternative assessment:

- Calls for complex reasoning, posing questions, making judgments, considering evidence
- Focuses on educational and communicatively worthwhile tasks, emphasising important learning goals
- Is based on multiple sources of data
- Recognises the learner as a unique person having multiple intelligences
- Is ongoing and emphasises learners' strengths and initiative
- Encourages collaboration between learners, sees classrooms as learning communities
- Assessment procedures and criteria are known to learners in advance, allowing choices and preparation
- Encourages teachers to develop meaningful curricula, providing detailed information to guide student learning
- Fosters intrinsic learning motivation

- Promotes responsible, self-regulated learning

The connection between alternative assessment and learning has previously been discussed in section 2.6.1.7, *Improvement in learning*).

The implementation of alternative assessment in recent years has required students to engage in an array of authentic tasks. It is clear from the general discussion of alternate assessment in Chapter 2 that alternative assessment provides a viable means of amending the limitations of traditional assessment (section 2.5.2, *Criticism of traditional assessment*).

Alternative assessment is vital in the language classroom as it provides students with the opportunity to use language in authentic settings. This type of assessment allows students to demonstrate their abilities using language in realistic contexts. An example of this would be writing a letter of application for a position in a summer camp or presenting a short piece introducing a favorite musician. Students are generally more motivated when they are presented with challenges and inspired to perform or present their work.

It is apparent from this thesis that a central aim of learning is to enable learners to be autonomous (section 2.2.3, *Constructivism*). Autonomous learning impacts EFL instruction and assessment since learners need to use language independently for authentic communicative purposes in order to attain their potential (Fisher, Evans, & Esch, 2004:51). Autonomy in language learning not only involves communication, but also allows learners to choose what topics they wish to study and how they want to learn (Macaro, 1997:170-171). Moreover, when pupils feel a connection to the learning material they become more involved in what they learn (section 2.2.3, *Constructivism*).

#### **4.12.1 Language for communication**

In order to teach language effectively, language should be taught for communicative purposes. Larsen-Freeman (2000:128-132) lists principles to facilitate communicative

language teaching, such as:

- The goal is to enable students to communicate in the target language.
- Teachers act as communication facilitators in the classroom.
- Almost all of the teaching-learning process is carried out with communicative intent.
- Teachers sometimes interact with students, but most of the time students interact with each other.
- The target language is used for communication and nonverbal nuances are also introduced.
- Language functions might be emphasised over forms.
- Teachers evaluate not only accuracy, but also fluency.
- Errors of form are seen as natural outcomes of the development of communicative skills.

If language teaching is aimed at communication, alternative assessment is essential as it provides the means wherein students can use communication in real situations. Alternative assessment activities allow an array of opportunities for the exchange of authentic language through communicative interaction.

#### **4.12.2 Content-based instruction**

Content-based instruction provides a natural context for natural users by integrating the learning of language with the learning of other subjects (Larsen-Freeman, 2000:137). The interest to assign meaning to language study has resulted in content-based instruction. Content-based instructions makes possible the natural integration of good language teaching procedures like alternative assessment, cooperative learning, integrated skills teaching, project work, strategy training etc. (Stoller, 2002:10). This method incorporates topics and tasks from many subject matters within the framework of teaching a second or foreign language (Crandall & Tucker, 1990:187).

The educational benefits of this approach are multifold.

- It is important for students to use languages for their own purpose and this can be achieved by allowing students to choose activities, materials and

conversational partners (Rigg, 1991:526; Stoller, 2002:9).

- Through content, learners are able to incorporate higher order thinking tasks (Met, 1991:282). Met (1991:150) suggest that: “content in content-based programs represent material that is cognitively engaging and demanding for the learner”. The activities chosen are intended to stimulate student thinking (Stoller, 2002:9).
- Learners can develop language skills while, at the same time, becoming more knowledgeable (Stoller, 2002:9).

Content-based learning has become a popular means of engaging language students in language and content learning (Stoller, 2002:9).

#### **4.12.3 Task-based learning**

Task-based learning has become a popular way of teaching foreign language. Littlewood (2004:319) sees task-based learning as a development and extension of communicative learning. Many writers seem to agree that there is a communicative nature to task work. Estaire and Zanon (1994:13-20) identify communication and enabling tasks as the two main categories of a task. Stern (1992: 195-196) mentions the connection between communication and task performance. Willis (1996:23) explains that tasks are activities in which language is used for a communicate purpose to achieve an outcome. Ellis (2000:195) concludes that many researchers and educators agree that communication is a vital element in task-based learning. It is apparent that task-based learning suits the communicative objectives of foreign language instruction.

The best way to assess task-based learning is through clear and explicit performance criteria. Criterion-referenced performance is beneficial in language learning as it focuses assessment on how language is mastered by matching student performance to clearly articulated performance standards. This, rather than on focusing on the performance alone (norm-referenced performance), can guide students towards more

successful language proficiency. The importance of criterion-referenced assessment in learning has been discussed (section, 2.3.3.7, *Criterion-referenced assessment*).

#### **4.12.4. Types of authentic assessment in first and foreign language learning**

There is a litany of methods that could be applied to assess both first and foreign languages. O' Malley and Valdez Pierce (1996:12) list the main types of authentic assessment in language learning:

- Oral interviews (of learners by the teacher)
- Story or text retelling (with listening or reading inputs)
- Writing samples (with a variety of topics and registers)
- Projects and exhibitions (presentation of a collaborative effort)
- Experiments and demonstrations (with oral or a written report)
- Constructed-response items (to open-ended questions)
- Teacher observation (of learners' work in class, making notes)
- Portfolios (focusing on learners progress over time)

According to O'Malley and Valdez Pierce (1996:12), an essential component of the different kinds of authentic assessment is that they: 1) concentrate on important curriculum goals (section 2.6.1.7, *Improvement of learning*), 2) aim at improving individual capability (section 2.6.1.2, *Continuous assessment and opportunities for improvement*), and 3) are conducted as an integral part of instruction (section 2.6.1.7, *Improvement of learning*).

Glaser and Silver (1994:26) maintains that assessment is closely linked to instruction and that the purpose of assessment is to assist in learning. Projects, performance tasks and portfolios are valuable assessment methods for the EFL classroom since they encourage formative assessment (section, 2.3.3.3, *Formative assessment* and criterion-referenced assessment (section, 2.3.3.7, *Criterion-referenced assessment*) which promotes more effective learning.

#### 4.12.4.1 Projects

Project work plays an important role in content-based instruction. It is useful as it provides a natural broadening of the content learning and language learning that takes place in classrooms (Stoller, 2002:10). Project work is valuable since it integrates the four necessary language skills - reading, writing, speaking and listening (Brown, 1994:218).

Stoller (2002:11) explains the many ways in which project work demonstrates the essential principles of language learning.

- Project work focuses on content learning rather than on particular language objectives. Authentic and interesting subject material can become the most important element of projects.
- Project work places the student at the center of the learning process. The teacher's role is to provide support and guidance.
- Project work emphasises the cooperative, rather than competitive nature of learning. Students have the opportunity to work alone, in pairs or in groups and are able to share resources, ideas and skills throughout the process.
- Project work allows for the genuine assimilation of skills and encourages the processing of information from a variety of sources.
- Project work concludes in a final product that can be shared with others. Since project work focuses on the ongoing process as well as the final product students have many chances to work on language fluency and accuracy at various phases of their work on the project.
- Project work can provide students with motivation and interest, and can challenge student thinking.

Stoller (2002:11) advises that a variety of factors, such as curricular objectives, course expectations, students' proficiency levels, students' interests, time constraints, and availability of materials must be considered when assigning project work. Although the presentation of the project is often the final stage, it is important to have students reflect on the whole project process. Stoller (2002:16) points out that by looking back

and evaluating their work, students gain a better understanding of what they have learned. For example, they can reflect on the language they used, the content knowledge they learned, the stages that they needed to follow to bring their work to completion and the merits of their final project.

Students who reflect on their work can effectively assess their own learning and gain a great deal of information on how they learn. They can recognise which stages of the process are difficult and which present particular challenges. This thesis has discussed the importance of self-assessment on learning (section 2.6.2.2, *Self-assessment*).

#### **4.12.4.2 Performance tasks**

Performance-based learning is the foundation of educational reform (Cody, 2000:15). This type of learning, which is based on the development and performance of authentic tasks, allows students to show what they know and can do with the knowledge they possess (Cody, 2000:15).

Williams and Burden (1997:168) define a task as “any activity that learners engage in to further the process of learning a language”. Willis (1996:23) explains that tasks are always activities where the target language is used by the learner for a communicative purpose in order to achieve an outcome. Stern (1992:195-196) emphasises the connection between tasks and realistic language use.

Skehan (2002:292) recognises the benefits that performance tasks have on language development. He suggests that students might recognise language forms they will need for a particular task and want to learn these forms in order to improve their performance. As a result, task-based learning can encourage and support language learning. O'Malley and Valdez Pierce (1996:5) agree that performance assessment tasks necessitate the integration of language skills which helps develop language.

Tasks have recently become a central pedagogical tool for language teachers. Psycholinguistic studies of second language acquisition and theories about learning foreign language propose that a learner's language develops through communicating

meaningfully through the target language (Williams & Burden, 1997:168). Therefore, performance tasks that require students to use skills and knowledge in the context of real life situations are especially suitable for ESL or EFL students who need to have experience using authentic language. Students can acquire foreign language by means of interacting, negotiating and transmitting meanings in the language in an effective way (Williams & Burden, 1997:168). They are a far better indicator of what students are able to do with language than traditional assessment tests.

Brown (2004:255) points out that it is important for teachers to relate to performances seriously and meticulously so that they can be valuable assessment instruments. In order to do this, he makes the following suggestions:

- State the general goal of the performance
- Specify the criteria of the performance in detail
- Prepare students for the performance by outlining the steps that must be followed
- Use dependable evaluation forms such as checklists and rubrics
- Relate to performances through feedback that is systematic
- If possible, use self and peer assessments prudently

The advantages of performance assessment benefit both teachers and students (Cody, 2000:17). Performance assessment provides teachers with:

- A deeper understanding of what students know and what should be done to further their learning
- A method of evaluation which is more interesting and relevant to students than other assessment methods
- A method of both discovering what students know and observing how they can apply the knowledge
- An efficient way to evaluate students in a standards-based educational system
- An adaptable method of assessing a variety of students' talents and abilities

Performance assessment provides students with:

- A more engaging and active way of learning
- More information about what they are expected to know and be able to do than

fixed-response tests, such as multiple-choice

All performance tasks require students to produce an actual product (such as an advertisement, an editorial, a photo journal etc.) or performance (such as a play, a video presentation, a speech etc.). These tasks must include specific instructions that explain what they are expected to do. A few possibilities are: persuade an audience, design a product, perform an experiment, solve a problem, interpret an event, analyse a perspective, compare and contrast, investigate an issue, provide support or make a decision (Cody, 2002:18). These activities require the use of higher level cognitive thinking skills which promote learning (section 2.2.2, *Cognitivism*). In addition, they are stimulating and challenging activities which can whet student interest and keep students interested and involved (2.6.1.4 *Authenticity*). Yet, in order to support performance tasks and assess them effectively, teachers need to be confident and adept in the area of performances (Stiggins, 2005:27).

#### **4.12.4.3 Portfolio**

Portfolios (section 2.6.2.1.2, *Portfolio*) are a useful tool for authentic assessment in the language classroom. They help learners develop language skills and encourage learners to become actively involved in the learning process (Nunes, 2004:328). Portfolios promote intrinsic motivation by transferring responsibility and ownership to the learner (Brown, 2004:257). The importance of motivation on learning has been discussed in section 4.5.3, *Motivational factor*.

Formative assessment has gained popularity in recent years for measuring student achievement (section 2.3.3.3, *Formative assessment*). Therefore, portfolios are valuable since they provide a continuous record of students' language development (Genese & Upshur, 1996:99). They encourage students to become involved with their own work. In addition, they have become popular with teachers who are interested in monitoring the learning process as well as learning outcomes (Smith, 1997:32). Nunes (2004:328) and Brown (2004:257) see portfolios as a vehicle that promotes continuing interaction between teacher and students.

Portfolios allow students to demonstrate what they can do with language in a more productive way than through traditional assessment (Tannenbaum, 1996). Since portfolios generally include a wide sampling of work, they allow students to exhibit much more than tests do (Hamp-Lyons, 1999:26). Brown (2004:257) agrees and credits portfolios for making it possible to assess multiple facets of language learning.

Furthermore, portfolios promote learning. In addition to the work that students must include, students are expected to revise and correct their work (Brown, 2004:257). Therefore, portfolios direct the students' attention to mistakes, and help them understand their strengths and weaknesses thereby increasing learner's responsibility (Smith, 1997:34). This offers motivation to improve, which accelerates and augments classroom learning (Smith, 1997:36). Portfolios offer a positive approach to assessment, as they are less threatening than tests.

Moya and O'Malley (1994), who advocate the use of portfolios in the ESL classroom, regard portfolios as an amalgam of several levels of knowledge, skills and capabilities. They maintain that a portfolio requires planning, design, implementation and evaluation. Portfolios, thereby, involve students in all stages of learning.

An integral part of portfolios should be a record of the learners' reflections on their work (Nunes, 2004:328). Looking back and reflecting on their own learning processes engages learners in cognitive thinking which is a vital element of learning (section 2.2.2, *Cognitivism*). Herbert (2001:55) agrees that reflective analysis improves learning. He insists that learners must be provided with the means to communicate and broaden their understanding of their own learning process. Reflections give students the opportunity to become better aware of their ability and progress. Reflections also provide the teacher with a wealth of knowledge regarding the learning methods, needs, and difficulties of their students. This knowledge enables teachers to adapt their teaching to meet and satisfy the needs of their students (Nunes, 2004:334).

#### **4.12.4.4 Self-assessment**

An interest in bringing the learner to the forefront of the learning process has resulted

in an interest in self-assessment (Alderson & Banerjee, 2001:227). The benefit of self-assessment is that it involves learners directly in the learning process and encourages greater autonomy in language learning (Brown & Hudson, 1999:666; Boud & Brew, 1995:9).

Mantle-Bromley (1995:383) maintains that if teachers address the affective and cognitive needs of students' attitudes, students will obligate themselves to more time in language study. The result will be evident in their academic accomplishments. When students have a personal interest in their learning they can get better marks (Harris, 1997:13). Self-assessment encourages students to become more involved in their own learning. Students become active participants in their own learning when they reflect on their learning and the outcome of their learning. Both these factors can increase motivation, which can facilitates language learning (section 4.5.3, *Motivational factor*) and encourage cognitive thinking (section 2.2.2, *Cognitivism*).

#### **4.12.4.5 Peer assessment**

There are clear benefits for encouraging students to reflect critically on the abilities of other learners who share a goal. First, considering someone else's work helps them assess their own performance. Second, the fear of making mistakes is eased when mistakes are recognised in other people's work (Tudor, 1996:182)

#### **4.12.4.6 Group assessment**

Group assessment is valuable in the EFL classroom because it promotes collaboration and self initiated language (Brown, 2001:177). Since the primary intention of language instruction is for communicative purposes (section, 4.12.1, *Language for communication*) group assessment maximises the opportunity for active language interaction. Kramsch (1992:17) agrees that the interaction of the social group is the vehicle through which language is used and learned. Group work facilitates language more effectively by encouraging fluency rather than accuracy (Brown, 2001:56).

Interaction within a group, rather than directly with the teacher, helps students feel more confident using language (Brown, 2001:178). Self-confidence is a decisive

factor in language acquisition (section 4.5.4, *Level of self-confidence*). Group assessment encourages students to be responsible and autonomous by causing them to focus their attention on a common objective (Brown, 2001:178). Students who take responsibility for their own learning outcomes are more effective learners (section 4.5.6, *Feelings of autonomy*).

It may, however, be that some children who can deal with teacher feedback feel threatened by their peers. Teachers, therefore, need to consider this possibility when they decide which kind of assessment to use.

#### **4.13 EFL TEACHING IN ISRAEL**

Teaching English in Israel presents several challenges. There are quite a number of dissimilarities between the English and Hebrew languages. The Hebrew alphabet, like other Semitic alphabets, is written differently than English. Although the sounds of most letters in the Hebrew alphabet have some correlation to existing sounds in English, their written representation has no apparent similarity. Additional differences exist in the linguistic syntax of the two languages. For example, Hebrew speakers are accustomed to the adjective succeeding the noun; for example, *habayit hagadol* is translated as *the big house*. When Hebrew speakers begin to learn English, they expect similar syntax. Therefore, it is difficult for them to remember to place the adjective before the noun, as they must do when using English.

A considerable challenge when Hebrew speakers learn English is that the direction of writing is from right to left and that some graphemes are visually identical, but differ auditorally. For example, 'o' in Hebrew sounds like 's' and 'l' sounds like 'v'.

An additional obstacle that Hebrew speakers face when they learn English is to understand tenses which do not exist in Hebrew. This has been discussed in section 4.10.4.2, *L1 supplies a foundation for language understanding*.

#### **4.14 ALTERNATIVE ASSESSMENT IN EFL IN ISRAEL**

As previously explained in this dissertation (section 3.10.1, *The New Curriculum*),

*The New Curriculum* introduced a change of thinking in regard to learning, teaching and assessment. Real learning began to be seen from a constructive viewpoint, where knowledge acquisition occurs over time. Alternative methods of assessment became needed to evaluate the new concepts in learning.

#### **4.14.1 Assessment Guidelines for the English Curriculum**

In order to provide teachers with the initial tools to begin to incorporate alternative assessment methods in the classroom, the Ministry of Educational Pedagogical Secretariat published the booklet, *Assessment Guidelines for the English Curriculum (2001)* in 2002. This booklet explains in detail the principles of classroom assessment that are outlined in the *English Curriculum* and will be referred to in this dissertation as *Assessment Guidelines*.

##### **4.14.1.1 Principles of classroom assessment**

The most important objective of assessment is to improve learning (section 2.6.1.7, *Improvement of learning*). This can best be achieved by creating and conducting assessment in the classroom where real learning occurs (section 2.6.1.8, *Takes place in the classroom*). In order for assessment to be an essential part of the teaching-learning process teaching objectives, teaching methods and assessment procedures need to be planned in advance and connected to each other. Teachers and pupils are encouraged to take part in all areas of this process (Ministry of Education, Culture and Sport, 2002:3).

##### **4.14.1.2 Categories of classroom assessment**

In order to give an accurate evaluation of a learner's progress, the *Assessment Guidelines* (2002:3) suggests using both formative and summative assessment in order to assess properly and accurately. Formative assessment, as explained earlier (section 2.3.3.3, *Formative assessment*), monitors a learner's progress and provides feedback that allows for continued and improved learning, whereas summative assessment (section, 2.3.3.4, *Summative assessment*) assesses the learner's

advancement of benchmarks and standards (Ministry of Education, Culture and Sport, 2002:3).

#### **4.14.1.3. Standards and benchmarks in assessment**

The *Assessment Guidelines* carefully explain the standards and benchmarks for each of the four domains of assessment so that teachers can evaluate their pupils' progress correctly. There are four distinct standards, each of which is accompanied by knowledge and performance benchmarks (Ministry of Education, Culture and Sport, 2002:11).

- **Standard: The Role of Assessment**

Teachers are aware of the role of assessment as an integral part of the teaching-learning process and assess the performance of their learners as part of their teaching routine.

#### **Knowledge Benchmarks:**

Teachers will meet this standard when they know about:

- The interdependency of teaching, learning, and assessment
- The importance of providing feedback and monitoring as essential for effective learning
- Methods of assessment that take into account different levels, learning styles, and abilities in heterogeneous classes
- The value of alternative assessment

#### **Performance Benchmarks:**

Teachers will meet the standard when they:

- Integrate teaching, learning and assessment in the planning of their unit lessons, and tasks
- Ensure that learners are regularly aware of their language learning process
- Design varied tasks and tools that allow learners to perform and succeed at different levels according to different learning styles and abilities

- Use alternatives in assessment including projects, portfolios, and presentations.

**Standard: The Methods of Assessment**

Teachers know about theories and methods of assessment and match them with the appropriate tasks and tools.

**Knowledge Benchmarks:**

Teachers will meet the standard when they know about:

- The distinction between formative and summative assessment
- The various types of assessment methods such as performance-based tasks, projects, portfolios, tests, and the goals they are designed to achieve
- A wide range of assessment tools such as rubrics, assessment lists, rating criteria, verbal and written feedback to evaluate learners' achievement of the different goals

**Performance Benchmarks:**

Teachers will meet this standard when they:

- Define for themselves and make clear to the learners the goals and criteria of the assessment task, prior to designing it
- Provide a varied range of assessment tasks
- Collect and record information about the learners' progress over a period of time from a variety of sources including homework assessment tasks, individual, pair and group activities

**Standard: The Learner's Role in Assessment**

Teachers are aware of the importance of involving learners and actively engaging them in the different stages of the assessment process.

**Knowledge Benchmarks:**

Teachers will meet the standards when they know:

- That learning is enhanced when learners feel ownership of the assessment procedures

- About different assessment tools that learners can generate and use to assess their learning such as peer and self-assessment
- About assessment tools that allow learners to evaluate both process and product of their performance

**Performance Benchmarks:**

Teachers will meet this standard when they:

- Encourage learners to contribute to the design of the assessment procedures such as determining criteria and writing test items
- Provide opportunities for learners to assess each other and themselves
- Encourage the use of assessment tools for learners to evaluate process and product, such as checklists and rubrics

**Standard: The Role of Testing in Assessment**

Teachers know about theories of language testing and design, and use tests appropriately.

**Knowledge Benchmarks:**

Teachers will meet the standard when they:

- Know about criteria for the design of tests (and other assessment methods) such as validity, reliability
- Know about the practical constraints in designing and administering tests
- Know about the appropriate ways of testing and assessing the skills and domains in the English Curriculum
- Know about a wide range of types of test items such as multiple-choice, open-ended, true/false, their advantages and disadvantages, and when it is appropriate to use them.
- Are aware of test anxiety and its implication
- Know about national test such as the Bagrut exams and their implications for teaching

**Performance Benchmarks:**

Teachers will meet the standards when they:

- Design tasks that are valid and reliable, and practical to administer and grade
- Include test items that are appropriate to the goals of the test
- Design tasks that have a balanced coverage of skills and domains as presented in the English Curriculum throughout the year
- Take steps to lower test anxiety by appropriate preparation, and by creating a supportive classroom climate
- Analyse test results using appropriate test calculations
- Take into account national tests in their teaching

The standards and benchmarks that were developed initiated the implementation of performance-based teaching and assessment. The *Assessment Guidelines* have been strongly influenced by the theories and principles of learning and assessment that were previously discussed (section 2.6.1, *Criteria for effective assessment* and 3.7, *Advantages of Standards*).

**4.14.1.4 Performance-based teaching and assessment**

The advantages of performance-based tasks on learning are clear (section 2.6.2.1.1, *Performance-based tasks*).

The *Assessment Guidelines* (Ministry of Education, Culture and Sport, 2002:12-13) supports performance-based teaching for several reasons. Performance based teaching:

- Enables pupils to use their knowledge and skills in realistic situations
- Measures the use of knowledge and skills in the context of practical tasks
- Focuses on process as well as the end product enabling pupils to solve problems and make decisions throughout the learning process
- Stimulates affective, social and metacognitive aspects of learning

#### 4.14.1.5 Characteristics of a performance task

The *Assessment Guidelines* (Ministry of Education, Culture and Sport, 2002:15) list the characteristics that teachers should consider when designing a performance task.

These characteristics are:

- It has various outcomes; it doesn't require one right answer.
- It is integrative, combining different skills.
- It encourages problem-solving and critical thinking.
- It encourages divergent thinking.
- It focuses on both the product and process.
- It promotes independent learning, involving planning, revising and summation.
- It builds on a pupil's prior experience.
- It can include opportunities for peer interaction and collaborative learning.
- It enables self-assessment and reflection.
- It is interesting, challenging, meaningful and authentic.
- It requires time to complete (adapted from Birbaum, 1997).

The factors mentioned above are in compliance with the tenets of effective assessment that have already been discussed (section, 2.2. *Theories Underlying Assessment Practices*).

### 4.15 MATRICULATION EXAMINATION – BAGRUT

The matriculation examination (Bagrut) has a very significant influence on foreign language learning and assessment in Israel.

#### 4.15.1 Background about Bagrut

The effect of standards on assessment in Israel is most evident in the changes that have been made in the matriculation exam (Bagrut). The changes reflect many of the principles that underlie standards-based education. In order to familiarise teachers with the new assessment changes, the Ministry of Education published the *New*

*Bagrut Assessment Handbook (NBA Handbook)* in 2003. This document will be referred to as the *NBA Handbook*.

The major change has been in the format of Bagrut. The three hour matriculation examination (worth 100 points) that existed prior to 2003 has been replaced with modular exams. The seven modules (A-G) are each one hour and fifteen minutes. They are divided into 1, 3, 4 and 5 point level examinations (Ministry of Education, Culture and Sport, 2003:1):

- a. The one-point examination consists of module A
- b. The three-point examination consists of module A, B and C. Each module is worth 33.3 percent of the grade for the three-point level.
- c. The four-point examination consists of module C, D and E. Each module is worth 33.3 percent of the grade for the four-point level.
- d. The five-point examination consists of module E, F and G. Each module is worth 33.3 percent of the grade for the five-point level

#### **4.15.2 Advantages of the new Bagrut**

There are many advantages in adopting the modular model (Ministry of Education, 2003:5, 10, 4).

- **Motivational aspect**

The Ministry of Education acknowledges the important role that motivation plays in the learning process. Therefore, the three hour matriculation examination that existed prior to 2003 has been replaced with modular exams. The purpose is to allow pupils to reach standards at different times. This is especially important for weak learners who might feel threatened by a long, demanding exam. The importance of making provisions for different learners has been explained (section 2.6.1.6, *Fairness*).

Pupils are tested on one module at a time, and when they progress to the next level, they will already have 33 percent of their final grade. This should encourage learners to strive to achieve more.

- **Pedagogical aspect**

Assessment mirrors the objectives of the curriculum. The new format of the Bagrut examination provides for testing all four domains on the written examination: social interaction, access to information, presentation and appreciation of literature, culture, and language.

- **Cognitive aspect**

The types of questions on the Bagrut require three different levels of comprehension:

1. Identifying explicit information
2. Identifying implicit information
3. Interpretive comprehension

The integration of advanced comprehension skills requires pupils to recognise, analyse and interpret information.

- In the past multiple-choice questions on the Bagrut examination were given the same number of points regardless of what they were testing. Now the points per item depend on the level of reading comprehension that is involved and not on the type of question that is being asked.

The change in the point system shows that higher level activities are awarded more value. This change was made in response to research that indicates the importance of test items that challenge students to engage in higher level thinking. The integration of higher level activities is in alignment with the theories of assessment that have been discussed (section 2.2, *Theories underlying assessment practices*).

### 4.15.3 Dissatisfaction with Bagrut

Although the New Bagrut was created after extensive research into language theories and principles, EFL teachers in Israel have expressed dissatisfaction with the new system.

In an online poll conducted by ETNI (English Teachers Network in Israel) from May 1<sup>st</sup> through May 31<sup>st</sup> 2005 on the subject of New Bagrut Assessment, three hundred and two teachers responded to the question:

The New Bagrut Assessment.....

1. is a good system or a better system than the old Bagrut and has been implemented well.
2. is a good system or a better system than the old Bagrut, but has been implemented well.
3. is better than the old Bagrut system, but still has many flaws, no matter how well it is implemented.
4. is not as good as the old Bagrut system.
5. is not a good system.

The results of the poll showed that only six percent of the respondents believe that the new Bagrut is a good system, thirteen percent believe that it is good or better, but has not been implemented well, thirty-three percent believe it is better, but still has many flaws, twelve percent believe it is not as good as the old Bagrut and thirty-seven percent believe that it is not a good system. The results clearly indicate that teachers find fault with the new Bagrut system.

An additional poll on the subject of the Oral Bagrut (which assesses the oral presentation of students' projects) conducted from April 11<sup>th</sup> through May 1<sup>st</sup> 2005 yielded quite different results. One hundred and twenty-nine teachers responded to the question:

Should the oral Bagrut be cancelled – yes or no?

Forty-three percent of the respondents voted yes

Fifty-seven percent of the respondents voted no

Teachers feel that the oral examination, which is intended to assess students' communicative abilities, is an important assessment tool. The Oral Bagrut will be discussed in section 4.15.5, *Performance Tasks in EFL in Israel*.

#### **4.15.4 Projects in EFL in Israel**

As the result of studies that document the success of performance-based language programs around the world, performance-based tasks have recently become a requirement in English language classrooms. In Israel, project work has become, not only a compulsory addition to the language curriculum, but is a mandatory component of the Bagrut for all high school students in the country.

The developers of *The New Curriculum* (section 3.10.1, *The New Curriculum*) were impressed by research on the advantages of performance-based learning and, therefore, have exerted tremendous resources in their campaign to make performance-based tasks a vital part of the classroom curriculum. They have created web sites, offered in-service training courses and set high standards for the project requirement for matriculation exams.

Projects are appropriate for learning and are encouraged for use in the EFL classroom for several reasons (Mann, Shemesh & Shlayer., 2002:18-19):

- Projects represent authentic activities.
- Pupil-chosen topics generate more interest than teacher-chosen activities.
- Projects promote student-student and student-teacher collaboration.
- Well-designed projects demand higher learning skills like problem solving or devising a research question.
- Projects can be presented to an audience and allow students to demonstrate their expertise.
- Projects are suitable to heterogeneous classes where weak and strong students work together in order to success.
- Projects can be created for all levels of language learners.
- Projects are open-ended and necessitate teacher/student collaboration to

design rubrics.

- Projects should contain a written part as well as an oral presentation.
- Projects help raise the standards of language learning and encourage areas of language learning that are not always handled in the traditional classroom setting.
- By utilising language in a natural way – not through repetition and drills or exercises - pupils learn to use and recognise language. Students develop writing skills by submitting drafts, rewriting and revising. Reading skills are enhanced when students read Internet, book, or magazine articles.
- Projects incorporate all four domains of language learning and help students attain the standards of each of the four domains of the curriculum:

*Social Interaction-* Students develop oral and written communication skills, enriching their vocabulary and improving their uses of accurate language.

*Access to Information-* Students are exposed to oral and written texts from a variety of sources while researching their chosen topic.

*Presentation-* Students have opportunities for presenting ideas and information on a variety of topics in varied formats in spoken and written English. There is opportunity for creative, multimedia oral presentation that augment written presentation.

*Appreciation of Literature, Culture, and Language* – Research and comprehensive learning about the topic enables students to acquire cultural, historical and social awareness.

Since the inception of performance-based assessment in Israel in the year 2003, every student who takes Bagrut in English must complete and present a project for inspection. In most cases this is done at the end of their English studies in the twelfth grade although capable students may submit their projects a year earlier.

Project work is a valuable component of language assessment as it takes place in the natural learning environment (section 2.6.1.8, *Takes place in the classroom*).

#### **4.15.4.1 Project requirements**

The requirements for the written and oral presentation of the project are outlined in The *NBA Handbook* (Ministry of Education, 2003: 21). They are:

- Cover page
- Table of contents
- Rationale for choice of topic
- Introduction
- Body
- Conclusions
- Bibliography
- Self assessment/checklists/reflection
- Optional components – pictures, maps, graphs, video, power point etc.

#### **4.15.4.2 Project aims**

Teachers must clearly define the aims of the project as well as the elements that must be considered to show that the objectives have been met. The following objectives have been set as guidelines for project work in Israel (Mann, et al., 2002:19).

Pupils will be able to:

- Learn English by exposure to authentic language, both written and oral.
- Write a paper in English using correct grammar, vocabulary, punctuation and spelling.
- Use resource materials by learning how to summarise main ideas, collated from a number of sources.
- Translate material from Hebrew to English if necessary
- Present their work orally, using visual and/or aural aides
- Design a research question and write a research paper using the standard conventions

- Become experts on the topic of their choice
- Work cooperatively as a group
- Work independently of the teacher
- Take responsibility for their part in the group
- Learn to negotiate their ideas and opinions and reach a consensus with others in the group
- Learn skills like surfing the Internet, using a data base, computer presentation software, and dictionaries

The objectives outlined above closely support the principles of effective assessment that have been documented (section 2.6.1.4, *Authenticity*; section 4.5.6, *Feelings of autonomy*; section 4.12.4.6, *Group assessment*).

#### **4.15.4.3 Stages of the project**

The *NBA Handbook* provides a summary of the different stages of doing project work in the classroom in order to guide learners towards the successful completion of the project. The steps given in the handbook (Ministry of Education, 2003:22) are:

- a) Choose a topic
- b) Ask a question that will help define the topic
- c) List possible sources of information
- d) Use sources
- e) Take notes
- f) Summarise
- g) Writes a draft and review the work with the help of a checklist
- h) Prepare the final written project
- i) Prepare and give an oral presentation

There are no specific guidelines that offer a suggested length for the project. Teachers are advised to consider the abilities of their students in regard to length.

#### 4.15.4.4 Project evaluation

The *NBA Handbook* stresses the importance of assessing the entire project process as well as the completed project. Therefore, pupils are given a final grade based on their working file and on the finished project or presentation. The working file should contain (Mann, et al., 2002:20):

- Drafts
- Resource material
- The completed weekly work schedule
- Notes of group discussions and decision making
- Drafts of interviews with sources
- Sketches and plans for the presentation

Assessing the working file is formative (it is continuous throughout the project) and summative (it gives a final grade for the finished project). Both facets of assessment are equally important as has been explained (section 2.3.3.3, *Formative assessment*; section 2.3.3.4; *Summative assessment*).

Projects are evaluated with a variety of pre-designated tools of assessment such as: checklists, rubrics, self evaluation, peer evaluation and verbal evaluation by the teacher (Mann, et al., 2002:20).

- **Rubrics**

Rubrics are a good way to evaluate student work. Students should be familiarised with rubrics and understand their format, what they are comprised of and how they are used (2.3.4.2, *Rubrics*).

Rubrics should be used in projects for several reasons (Mann, et al., 2002:28):

- They let assessment be objective and reliable.
- They show pupils how their work will be evaluated and what is expected.
- They make pupils aware of the criteria used in assessing their own and peer performance.

- They provide effective feedback to pupils about their strengths and weaknesses and how they can improve.
- They provide benchmarks and standards of proficiency against which the pupils' performance will be measured.

The *NBA Handbook* provides rubrics for assessing project work (Ministry of Education, Culture and Sport, 2003:25). The pupil can receive a maximum number of 80 points (an additional 20 points are awarded for the oral presentation of the project) which are divided in the following way:

Content: 0-25 points

Sources: 0-10 points

Language: 0-15 points

Effort: 0-10 points

Individual Contribution: 0-20

Students need to use their skills and knowledge to gather information, analyse and produce the content of their project. Students are encouraged to take ownership of their work in order to make their individual contribution to the project.

#### **4.15.4.5 Reflection and feedback**

The final stage of the project is pupil self-reflection and feedback from the teacher. It is suggested that pupils relate to several of the following questions when reflecting on their work (Mann, et al., 2002:25):

- Did you enjoy working on the project?
- Did you enjoy working in a group or pairs?
- What advantages do you think group work offers?
- What are the disadvantages of group or pair work?
- What new words did you learn?
- What new skills did you learn?
- Which part of the project did you enjoy the most: the research, the writing, the

oral presentation?

- Which part of the project was the most difficult for you and why?
- Which part of the project was the easiest for you and why?
- Next time you do a project, what changes would you make?
- In what area do you think your English has improved: reading, writing, spelling, new vocabulary, grammar, speaking?
- In what ways do you think your learning strategies have improved?
  - I have become more independent.
  - I am more organized.
  - My group can rely on me to do my share.
  - I have more confidence in my abilities.
  - I know I can get help if I need it.
  - I have learned how to work in a group.
  - I am no longer afraid of difficult texts.
  - I know how to use a dictionary for help.
  - I understand the importance of writing a first draft.
  - I have learned to be critical about my work.

Teachers are instructed to give feedback supplied by rubrics. These rubrics should consist of three features (Mann, et al., 2002:23):

3. It shows the criteria according to which the assessment will be made.
4. It describes the standards of proficiency for each criterion.
5. It reflects various levels of proficiency and the weight of each level in a grade.

In addition to the feedback given by rubrics, teachers are instructed to provide written comments. These comments have three parts (Mann, et.al, 2002:25):

2. A positive comment – encouraging words- that reflect pupils' effort collaboration, achievements, presentation and understanding of the topic.
3. Constructive comments in simple language that identifies areas that need additional work.
4. Suggestions for improvement, enrichment and advancement:
  - You should read more.
  - Spend a few minutes every day practicing your spelling.

- Make a personal vocabulary list for yourself and learn the words.
- Practice speaking aloud to a friend or tape yourself and listen to the tape.
- Read an article from an English language newspaper every day.

The written project must be followed by a presentation. Students are taught how to present their project and focus on the important elements of their written work. Project and presentation work are considered to be important elements of language study. As a result, both have been included as required components of Bagrut. The oral presentation based on the written project comprises twenty percent of the total Bagrut grade. The importance of performance tasks is discussed in the next section.

#### **4.15.5 Performance tasks in EFL in Israel**

Pupils are encouraged to use a variety of presentation formats in their oral presentation, such as: showing a video, presenting a PowerPoint presentation, creating a website, using visual aides and incorporating music ((Ministry of Education, 2003:22). The influence of multiple intelligence theory is clearly evident in the planning of both projects and performance tasks (section 2.2.4, *Multiple intelligences*).

The *NBA Handbook*, (Ministry of Education, 2003:28) outlines the several benefits of performance tasks:

- Builds background knowledge
- Has significance for the pupil because it enables problem-solving
- Enables multiple responses and learning outcomes
- Enables collaborative learning
- Focuses on multiple dimensions of pupils' learning
- Allows pupils to select a presentation mode to display what they have learned
- Requires pupils to practice, refine and revise in order to demonstrate their learning
- Enables self-assessment and reflection

Performance tasks encompass modern educational practices and theories. Firstly, they are in harmony with cognitive theories that acknowledge the importance of students drawing on their own experiences and understanding, as well as solving complex problems (section 2.2.2, *Cognitivism*). Secondly, these tasks encourage a variety of performances, styles and individual expressions, which are features of the theory of multiple intelligences (sections 2.2.4, *Multiple intelligences*). Thirdly, performance-based tasks allow for group interaction which benefits both the individual students and the group (section 2.6.2.3, *Peer assessment and group assessment*). Lastly, these tasks provide motivation and nurture feelings of self-confidence by allowing students to review, refine and revise their work.

#### **4.15.6 Portfolio in EFL in Israel**

Portfolios are widely used in EFL classrooms in Israel. They are considered to be a valuable learning tool for several important reasons. For example, portfolios encourage independent and active learning, they develop student awareness of their own progress, they allow for the assessment of a variety of skills and they accommodate individual learning styles and interests (Kemp & Toperoff, 1998:2). Portfolios provide students with a visual illustration of their language abilities (Shlayer, 2000: 37).

The use of portfolios in teaching has had a major impact on learning and assessment in Israel (Smith, 1997:33-36).

- **Portfolio and learning**

Students working on portfolios need to be diligent about preparing work on time and according to set requirements. Since they are required to correct and revise assignments they pay attention to details and learn from their mistakes. By applying time and attention to revision they gain a better understanding of their strengths and weaknesses in their language study. This understanding is a form of cognitive awareness which is accelerates effective learning (section 2.2.2, *Cognitivism*).

- Portfolio and assessment and testing

Assessment is made according to pre-decided criteria, known to the learners and written on a separate sheet appended to the portfolio. The premise is that if students are aware of the criteria that are considered in the portfolio grade, they will expend time and effort in these areas. For example, if points are given for revision and for reflection, students will be careful to do these things. Students gain a better understanding of their own abilities through portfolio assignments, revision and reflection. They can feel a sense of accomplishment and are rewarded for the effort they put into the learning process, not only for the outcome. Whereas in traditional assessment tests measure the outcome of learning, portfolios assess the learning process (section 2.6.1.7, *Improvement of learning*).

- **Guidelines for portfolios**

The *Guidelines for Portfolio Assessment* (Kemp & Toperoff, 1998) is the official guideline document approved by the Ministry of Education for portfolio work in Israel. The document explains:

- The purpose of a portfolio
- Reasons for using portfolio assessment
- Essential elements of a portfolio
- Implementing portfolio assessment
- Assessing the skills through portfolio

Portfolios have important implications for assessment (Kemp & Toperoff, 1998)

- Portfolios items are products that have been produced in class during the natural learning process. They are not disconnected from everyday learning like tests may be.
- Portfolios have clear goals that are decided upon by the student and the teacher at the beginning of instruction.
- Portfolios give a profile of learner abilities. They show students' ability to use a variety of skills, work with resources, and improve and develop over time.
- Portfolios allow students to work at their own level and according to their

learning style and interests.

- Portfolios develop independency – students choose their portfolio choices and monitor their progress.
- Portfolios show learning progress over time.
- Portfolios foster teacher-student dialogues.

#### **4.16 THE CHALLENGES OF ALTERNATIVE ASSESSMENT IN EFL IN ISRAEL**

The developments in EFL teaching have been documented in this chapter. The Ministry of Education in Israel has set into motion widespread reforms that affect the way language is taught and assessed in the country. However, despite the enthusiasm of many teachers, there are a considerable number of teachers who are reluctant to use alternative assessment. As a result, there are many difficulties in implementing alternative assessment in foreign language teaching in Israel.

- **Not directly related to the Bagrut**

The Bagrut is a set format. Therefore, teachers (particularly in high school) may feel that using alternative assessment does not adequately prepare students for the type of assessment they will meet (Maoz, 2000:91).

External tests have a strong influence on the content and format of classroom tests (Shohamy, 2000:58). Therefore, teachers imitate and duplicate external tests in classroom testing because they believe that these tests are better. Bagrut is based essentially on traditional assessment test items.

- **A change in attitude**

Many teachers are not familiar with alternative assessment. They have relied heavily on tests. Teachers who have taught for years may have to change their attitude regarding their teaching and approach to assessment (Shlayer, 2000:38).

Although teachers have begun to adopt alternative assessment in their classes, they still rely primarily on tests rather than students' performance when they assign final grades (Shohamy, 2000:54).

- **Training**

There is a need for ongoing training in-service training to help teachers understand and implement alternative assessment (Shlayer, 2000:38). Understanding and implementing the new curriculum depends on personal and professional change.

Teachers may be reluctant to use alternative assessment because they lack the knowledge and training that is necessary to incorporate alternative assessment in the classroom (Shohamy, 2000:58).

Some teachers, who receive training, may feel that the training is not directly related to their classroom teaching and is not practical (Shohamy, 2000:58). However, teacher trainers who are aware of the potential problems involved in the process of change can provide teachers with the necessary tools to facilitate professional development (Rechtman, 1999: 23).

- **Time consuming**

A negative aspect of alternative assessment, such as project and portfolio work is that it is very time consuming for teachers, and demands a considerable amount of extra work (Shlayer, 2000:38; Maoz, 2000:91)

Projects that are done outside the classroom might not be authentic work, while permitting tasks to be done only during lessons occupy too much class time (Maoz, 2000:91).

- **Fear of the unknown**

An additional obstacle to alternative assessment, primarily project work is the

reluctance and apprehension of many teachers to leave the safe haven of their classrooms to face the technical challenges of the computer laboratories.

Project work demands that teachers assist their students with the skills required to complete extensive research projects (Ministry of Education, Culture and Sport, 2003:21-27). Many language teachers have not been trained to work with computers, and lack the skills that are necessary to guide their students. As a result, they lack confidence in their abilities and may not be enthusiastic about project work.

Choice plays an important role in alternative assessment. Teachers in Israel are uneasy about encouraging their students to make choices since it is not a fundamental part of teaching methodology in the country (Maoz, 2000:91).

In addition, teachers must grapple with discipline problems, which can result in the computer laboratory where lessons are less structured. Despite the importance of individualised learning which is clearly an advantage of project work, many teachers who are unaccustomed to this mode of instruction become uncertain and uncomfortable.

#### **4.17 SUMMARY**

The development of foreign language teaching and learning in Israel has been influenced by many factors. Cognitive theories, which shed light on how students learn, have been incorporated into EFL teaching to facilitate language acquisition. Classroom activities including alternative assessment, such as projects, presentations and portfolios challenge students to use higher level thinking and help them become autonomous learners. Self-evaluation and reflection are considered to be essential elements for productive learning.

In addition, the standards movement in education, which extended to include standards in foreign language study, has had a dramatic influence on assessment. This impact is evident in greater attention to criterion-referenced assessment and formative assessment, both of which have been shown to motivate students and assist learning.

The shift in focus to language for communicative purposes has effected language instruction and assessment. This is apparent in the role of globalised education in the classroom where assessment centers on the use of authentic material and communicative language.

Research into language study has changed the approach to language assessment in Israel. Matriculation examinations have been revised so that they reflect the importance of assessing higher level thinking. Many areas of classroom assessment have become more rigorous and challenging for students.

The next chapter will discuss the research and design methodology of the empirical study.

## **CHAPTER 5**

### **RESEARCH DESIGN AND METHODOLOGY**

#### **5.1 INTRODUCTION**

This chapter presents a description of the research plan that was employed in the empirical study. It outlines the procedures that the researcher used to collect and process data that was received in response to the research questions that were introduced in the study. Section 5.2 describes the objectives of the empirical study. Section 5.3 details the research methodology utilised in the study and explains the decision to adopt a qualitative research design by presenting the advantages, disadvantages and justification for using this type of research. Section 5.4 deals with sampling and explains why a particular sampling was used in the study and gives the rationale for the dependability of the chosen sample. Section 5.5 discusses the role of the researcher as an instrument in the collection of data in this study. Section 5.6 explains the procedures that were used to collect data. Section 5.7 details the various methods that were applied to collect the data such as textual analysis and focus group interviews. This section provides an understanding of these methods, their functions, their advantages and their limitations as well as a justification for relying on their accuracy. Section 5.8 discusses the role of field notes and field analysis. Section 5.9 explains how the data was compiled, checked, coded and analysed in order to present relationships among the many variables. The final section, section 5.10, presents a summary of the chapter.

#### **5.2 OBJECTIVES OF THE EMPIRICAL STUDY**

The purpose of the empirical study is to indicate how EFL secondary school teachers in Israel have implemented the principles of alternative assessment as was discussed in the literature study. The empirical study explores the degree to which teachers are aware of the theories and principles of assessment and to what extent they have implemented these principles into a standards-based curriculum and the subsequent alternative assessment practices required.

The study begins with the following set of hypotheses:

**Hypothesis 1**

Assessment in EFL in Israel still heavily relies on traditional assessment methods mainly because teachers are reluctant to use alternative assessment

**Hypothesis 2**

The Bagrut has a significant influence on assessment by causing teachers to choose assessment that reflects Bagrut assessment

**Hypothesis 3**

The modularised format of the Bagrut has resulted in non-integrated teaching of language skills.

**Hypothesis 4**

Teachers are unfamiliar with the principles of standards and the relationship between standards and assessment.

**Hypothesis 5**

Teachers don't actively involve their students in classroom assessment.

### **5.3 RESEARCH METHODOLOGY**

The research methodology in this study utilises several qualitative research designs in order to show how EFL teachers in Israel relate to a standards-based curriculum and how they integrate alternative assessment into their teaching. An understanding of teachers' beliefs and practices was gleaned from data collected from textual analysis and focus group interviews. The textual analysis was comprised of a study of the assessment packages of ten 11<sup>th</sup> grade EFL teachers. The assessment packages were evaluated according to the principles of assessment practices explained in the literature research chapters of this thesis (Chapters 2, 3 and 4).

In addition to the textual analysis, focus group interviews were conducted in order to determine teachers' practices regarding the implementation of a standards-based

curriculum and to establish how they implement standards and alternative assessment principles into their lessons.

The intention of this study is to enrich the understanding of the current trend in EFL teaching and assessment.

### **5.3.1 Qualitative research: a theoretical basis**

Over the years qualitative research has gained popularity in educational investigation because of the advantages of research by means of direct contact (Johnson, 1995:4). Since the focus of this study is an investigation into the actual practices of teachers, it warrants direct interaction with teachers in the field. Qualitative research is the method that provides the best opportunities to gather data through observation, interviews and field data records (McDonough & McDonough, 1997:53). It offers the researcher the opportunity to obtain information directly from the people whose opinions and practices are the subject of the study.

The researcher will discuss the advantages and disadvantages of qualitative research. The rationale for choosing this method as well as the justification for accepting the reliability of the results will be explained.

#### **5.3.1.1 Characteristics of qualitative research**

There are several factors that define qualitative research (Grady, 1998:5-10) and distinguish it from quantitative research.

- The purpose of qualitative research is to understand the reasons for certain situations (unlike quantitative research whose purpose is prediction and control).
- The focus of qualitative research is on a complete and holistic understanding of the subject (unlike quantitative research which studies narrow or selected variables).

- Qualitative data are usually words and phrases and statements, which the researcher examines in order to establish themes that respond to the research questions (unlike quantitative data which are usually numbers used for statistical purposes).
- Qualitative instrumentation employs a researcher as the main data-collection instrument (quantitative data-collection techniques are usually tests and questionnaires).
- Qualitative research is usually carried out in natural conditions (quantitative research is usually carried out under controlled situations).

### **5.3.1.2 Advantages of qualitative research**

This empirical study uses qualitative research methodology because of the many advantages that such a study provides. Some of the advantages of a qualitative research study are:

- It allows for greater flexibility in a less structured and controlled setting (Patton, 1990:55; Creswell, 1998:15; Grady, 1998:4).

The fact that qualitative research allows for more flexibility was beneficial to this researcher because the participants could themselves agree upon the time for the focus group interviews.

- It provides the opportunity for direct contact with subjects in a group format (Vaughn, Shay & Sinagub, 1996:12).

The participants in this study were willing to express their thoughts and views and interact with other members of the group.

- Its reports are descriptive and include expressive language. This results in richer information that allows for wider interpretations and deeper

understandings than statistical and numerical responses (Eisner, 1991b:36).

The participants expressed themselves well and offered examples and explanations.

- The researcher can construct a composite, holistic picture, analyse words and report comprehensive views of respondents rather than depend on statistics (Creswell, 1998:15).

The comprehensive discussions resulted in a full understanding of the subject.

- The rich language and expressive descriptions allow for more open and varied interpretations of the data (Grady, 1998:8).

The choice of vocabulary and language structure contributed to the content of the discussions. Participants' tone of voice and language use gave an indication of their interest and enthusiasm for the subject.

- The researcher assumes a prominent role in qualitative research. Assumptions and conclusions are not reached from the gleanings of statistical data, but rather from the keen understanding, and sensitive analysis reached by the researcher (Strauss & Corbin, 1990:42).

The researcher's familiarity with the subject of the study enabled her to understand, analyse and interpret the information gathered during the group discussions.

Although there is a litany of reasons to support qualitative research, adversaries of this practice have defended their opposition.

### **5.3.1.3 Criticism of qualitative research**

One criticism that has been directed at a qualitative approach to research is that it does not have the credibility connected with traditionally accepted quantitative methods in which investigation relies on numerical measurement (Horsburgh, 2002:308).

An additional criticism of qualitative research is that it lacks objectivity (Walker & Evers, 1999:43). Therefore, the results of qualitative research can be problematic because of the researcher's personal involvement in the research and his or her personal history, experience and perspectives, which might influence the analysis (Grady, 1998:7).

The researcher enlisted the following measures to circumvent the disadvantages presented by a qualitative study:

- The focus group was comprised of a sampling of teachers who represented various teaching populations, for example males and females; and teachers who taught higher and lower level classes. They were also teachers with varying degrees of experience, with varying degrees of professionalism, those with and without children of their own who learned English in Israel or abroad.
- The researcher acted as a facilitator and was careful not to project personal views either during the discussions or when collating and analysing data.
- Triangulation was achieved through the evaluation of data from textual analysis and focus group interviews. The use of triangulation (refer to section 5.3.2, *Triangulation*) was utilised to ensure a more reliable research outcome than a singular method. In addition, since some of the data collected could be interpreted numerically this allows for a quantitative variable.

#### **5.3.1.4 Justification for using a qualitative research design**

The type of methods used in educational research should be based on how well they can generate useful information for testing a hypothesis (Mayer, 2005:76). A qualitative research design was adopted to allow the researcher to interpret the significance and perspective of the subject being researched (Popay, Rogers & Williams, 1998:347).

This particular study can best give information and provide insight through spontaneous and open discussion. Qualitative research recognises that people are the best qualified to describe, interpret and discuss their own environment (Anderson & Arsenault, 1998:134). In addition, the natural and comfortable setting can encourage subjects to respond candidly, which adds meaning to any study.

Although qualitative design methods have been criticised for being subjective, no research is completely free of prejudices, suppositions and the personality of the researcher who is closely engaged in the study (Sword, 1999:277). Researchers who have a close affinity to their subject may project personal feelings into their research design and execution whether it is in the collection of data or its analysis.

In order to assure validity, the researcher adhered to the guidelines outlined by Anderson and Arsenault (1998:134), and kept meticulous records of all sources of information, used detailed transcripts, took field notes of all communications and engaged in reflective thinking during the research.

### **5.3.2 Triangulation**

Triangulation, which means using a number of types of data collection to approach the same issue or question, is an essential element in qualitative research (Grady, 1998:31). In qualitative studies triangulation most often means utilising data collection techniques, such as interviews, observations and document analysis in which each method substantiates the data from the other (Grady, 1998:31). Using a number of data-collection methods helps to ensure the reliability of the information that is gathered by avoiding bias or the distortion of the researcher's representation of the findings (Grady, 1998:31; Anderson & Arsenault, 1998:131).

The benefit of triangulation is that when two or more sets of findings are in agreement, there is greater certainty regarding the results of an investigation and greater confidence in the reliability of the data (Cohen & Morrison, 2000:112). Although Grady (1998:31) maintains that three or more methods are needed for triangulation, Cohen and Morrison (2000:112) maintain that two methods suffice, but that more can be used if deemed necessary.

The triangulation used in this study was a textual analysis and focus group interviews.

#### 5.4 SAMPLING

There are several types of sampling. *Convenience sampling* refers to individuals who are selected particularly because they are willing and available to be studied (Johnson & Christensen, 2004:214). This type of sampling can provide useful information, although the researcher cannot confidently state that the individuals chosen are representative of the population (Creswell, 2005:149).

*Probability sampling* refers to individuals who are selected from the population and are representative of that population. The researcher can claim that the sample is representative of the population and, as such, can make generalisations (Creswell, 2005:146).

*Purposeful sampling* refers to a population who has the specific characteristics that the researcher is interested in studying. The ability to generalise the results from this type of sample to the general population is limited (Johnson & Christensen, 2004: 215).

Sampling in a qualitative study should concentrate on the relevance of the sampling rather than on achieving a random and representative section of the population (Popay et al., 1998:346; Horsburgh, 2003:311). This is because the objective of a qualitative study is to garner rich and comprehensive information, which can best be provided by subjects who have the ability to supply relevant data on the subject under investigation.

Patton (1990:184) points out that sample size depends on the objectives of the study, and the acquisition of information that can be useful and reliable with the time and resources that are available.

The study was comprised of a convenience sampling. The intention of the sampling was to select subjects who are experienced and knowledgeable in the subject areas and who were agreeable to participating in the research study. The particular teachers

who participated in the focus group interviews, and who submitted their assessment packages and received the questionnaires were chosen for their availability and willingness to be part of the empirical study. The participants are all currently engaged in EFL teaching in high schools in Israel.

The subjects were comprised of both male and female teachers whose years of teaching experience varied. All the participants were familiar with the new requirements set by the English Inspectorate of the Ministry of Education in Israel and each participant had participated in at least one in-service training course on the new curriculum.

It was apparent that the success of the focus group interviews would depend on the ability of the group members to interact freely and honestly. Therefore, it was expected that the particular teachers who were selected would articulate their thoughts and actively engage in a group discussion.

For the textual analysis the assessment packages of ten 11<sup>th</sup> grade EFL teachers were analysed. Teachers were approached in a random way and the first ten teachers who were willing to submit their material for inspection and analysis were selected. Four of the teachers who were approached refused to participate, citing heavy work commitments as their reason.

The participants of the focus groups were comprised of several teachers, four of whom had submitted assessment packages. Twelve teachers agreed to participate in the discussions. Two focus groups were held with six participants in each focus group.

The sampling for the focus group interviews thus included only teachers who were knowledgeable and experienced in the subject area and who met the following criteria:

- They currently teach 11<sup>th</sup> grade EFL classes under the auspices of the Ministry of Education in Israel. The reason for the limitation was to ensure that the

teachers who participated in the study were aware of the recent changes in EFL teaching in Israel and have begun to implement these changes into their teaching and assessment practices.

- They were aware of the requirement set by the English Inspectorate of the Ministry of Education in Israel to implement a standards-based curriculum and alternative assessment tasks into their teaching.
- They completed, at least, one in-service training session on the new curriculum.
- They were willing to participate in a focus group interview.
- The researcher believed that they would feel comfortable and willing to share their thoughts, ideas and feelings.

The teachers who were requested to submit assessment packages were chosen for the following reasons:

- They currently teach 11<sup>th</sup> grade EFL classes under the auspices of the Ministry of Education in Israel. The reason for the limitation was to ensure that the teachers who participated in the study were aware of the recent changes in EFL teaching in Israel and have begun to implement these changes into their teaching and assessment practices.
- They were aware of the requirement set by the English Inspectorate of the Ministry of Education in Israel to implement a standards-based curriculum and alternative assessment tasks into their teaching.
- They were willing to submit their assessment packages for review in this study.

## **5.5 THE ROLE OF THE RESEARCHER**

The researcher is invaluable in any qualitative research study as he or she contributes a great deal to the success of the study. An experienced and qualified researcher can build rapport, which will help elicit rich data since subjects encouraged by good rapport are willing to contribute more information to the study (Grady, 1998:8). In addition, the quality of the research is highly dependent on the skills of the researcher to understand, record, gain insight and interpret the data collected (Anderson & Arsenault, 1998:134).

The researcher must have the skills to direct the interview without personal bias, which could manifest itself in leading questions or comments (Grady, 1998:20). Moreover, the researcher must make every effort to ensure confidentiality so that the interviewees do not feel inhibited or restrained (Grady, 1998:20).

The researcher needs to have the necessary skills to control the group, involve any silent participants into the discussion and restrain those who monopolise the discussion (Anderson & Arsenault, 1998:204).

I approached the role as researcher with an understanding of what is necessary in order to elicit involvement from the participants. I was aware that if the teachers felt comfortable they would be encouraged to express their views more freely. I prepared my questions and practiced my delivery several times before the interview so that I would be able to listen and think about the responses simultaneously. I reminded myself that it was important to pay attention to emotional responses as well as verbal ones.

## **5.6 DATA COLLECTION**

The process of data collection for the focus group interviews involved taking careful notes of the interviews, audio-taping the sessions, meticulously transcribing the texts of the audio-tapes and reading, organising and coding the data. A coding system was created based on key words within the data. Careful attention was directed to discarding irrelevant information recorded during the interviews.

The method of data collection for textual analysis consisted of an assiduous reading and analysis of the assessment packages. The data was organised according to the criteria contained in the research questions and then further systematised into sub-categories by using a coding system based on key words. Careful attention was directed to discarding irrelevant information recorded during the interviews.

The criteria used will be discussed in sections 5.7.1.1 (*Criteria used to examine and judge the influence of standards on assessment*), and 5.7.1.2 (*Criteria used to measure whether alternative assessment measures were incorporated*).

## **5.7 METHODS**

The researcher used the following methods of data collection: textual analysis and focus group interviews.

### **5.7.1 Textual analysis**

Textual analysis (or document or content analysis) is a research method that utilises a collection of procedures to formulate valid inferences from a text (Weber, 1990:9). Textual analysis is a major data collection technique (Grady, 1998:24) which is useful in identifying or solving educational problems (Borg & Gall, 1989:520). The purpose of textual analysis is to identify appropriate categories and units of analysis, both of which will reflect the nature of the document being analysed and the purpose of the research (Cohen & Morrison, 2000:164). It is valuable as it elucidates the content material as well as the source of the communication, and those to whom the information is directed (Cohen & Morrison, 2000:165).

For textual analysis to be effective there must be a clear and direct connection between the text and content that is studied and the objective of the research (Borg & Gall, 1989:523). The assessment packages of ten EFL high school teachers were collected and analysed to ascertain how closely they adhered to the principles and theories of standards-based education and alternative assessment.

In order for the textual analysis to be reliable, the sampling must be representative of the population from which it is obtained (Cohen & Morrison, 2000:95). Therefore, this study centered on a textual analysis of the assessment packages used by ten EFL high school teachers in their 11<sup>th</sup> grade English class in Israel. The researcher decided to focus on 11<sup>th</sup> grade classes because students generally study one module at this level which allows for more opportunity for diverse classroom assessment. Each teacher submitted the quizzes, tests, exercises, assignments and criteria that he or she used to assess the students over a semester period (five months).

The material submitted was studied and analysed diligently by means of criteria according to the principles of standards and assessment that were presented in the theoretical chapters (Chapter 2, *From traditional assessment to alternative assessment*, Chapter 3, *The implementation of standards in education and its impact on assessment practices* and Chapter 4, *Factors influencing the teaching and assessment of English as a foreign language in Israel*) of the thesis. The specific criteria are listed in the following section, 5.7.1.1, *Criteria used to examine and judge the influence of standards on assessment*. In addition, each assessment package was evaluated to ascertain to what extent teachers incorporated the standards outlined in the new curriculum (Section 3.10, *Standards in Israel*). Special attention was directed to the criteria used for assessment and its connection to the principles underlying the standards.

Two sets of criteria were used. One set was used to examine and judge how standards influence classroom assessment. The other set of criteria checked the extent to which alternative assessment played a role in classroom assessment.

The focus of text analysis is less on the units of the text, but rather on the issues of relevance to a particular assertion or debate (Krippendorff, 2004a:788). This study focuses on the degree to which the assessment practices of teachers comply with standards and assessment principles.

### 5.7.1.1 Criteria used to examine and judge the influence of standards on assessment

The central concept in analysing and judging a text is to establish the criteria that will be used as a form of reference in the analysis. These criteria originate from the research question and the theoretical study (Mayring, 2000).

Specific criteria were used to examine and judge how well teachers' assessment practices are aligned with the objectives and directives of the new standards-based curriculum in the teaching of EFL in Israel high schools. These criteria were based on the research material discussed in the theoretical chapters of this paper.

- The tasks demonstrate what students are required to know and be able to do.  
(Refer to sections 3.2, *Definition of standards*; 3.5, *Kinds of standards*.)
- The task is an exemplary demonstration worthy of emulation.  
(Refer to section 2.6.1.7, *Improvement of learning*)
- The tasks evidence the vital basic knowledge in particular subject areas rather than rote learning.  
(Refer to sections 3.2, *Definition of standards*; 3.5, *Kinds of standards*)
- The task encourages cognitive growth.  
(Refer to section 2.2.2, *Cognitivism*)
- The task maximises students' learning potential.  
(Refer to sections 2.3.1, *Assessment and evaluation*; 2.3.2, *Testing and assessment*)
- The tasks are specific and clear and can be easily assessed.  
(Refer to section 3.4.3, *Standards must be teachable*)
- The tasks require students to use their knowledge and skills in authentic situations.  
(Refer to sections 2.2.4, *Multiple intelligence*; 3.5, *Kinds of standards*)

- The task requirements are flexible, allowing for student diversity.  
(Refer to section 3.6.6, *Fails to consider the needs of diverse learners*)

- The tasks allow for students to utilise their multiple intelligences.  
(Refer to section 2.2.4, *Multiple intelligence*)

### **5.7.1.2 Criteria used to determine whether alternative assessment measures were incorporated**

Assessment methods were considered, analysed and evaluated to determine how well they reflect the findings of the theoretical research, with particular attention to whether alternative assessment methods were incorporated (Chapter 2, *From traditional assessment to alternative assessment*).

The assessment methods implemented by the ten teachers in the study were considered according to the following criteria:

- Traditional questions, such as: multiple-choice, matching, true/false, fill-in, essays and paragraphs.  
(Refer to section 2.2.1, *Behaviorism*)

- The memorisation and rote learning of isolated bits of information are required and encouraged.  
(Refer to sections 2.2.1, *Behaviorism*; 2.5.1, *Types of traditional assessment*)

- Students are assessed on social interactions in authentic settings and situations.  
(Refer to sections 2.2.2, *Cognitivism*; 2.5, *Traditional assessment*; 2.6, *Alternative assessment*)

- Students are actively involved and engaged in the assessment process.  
(Refer to section 2.6, *Alternative assessment*)

- Opportunities for self-assessment and peer assessment are provided.  
(Refer to sections 2.2.3, *Constructivism*; 2.5, *Traditional assessment*; 2.6, *Alternative*

*assessment*)

- Assessment encourages varied mediums for assessment, such as: journals, projects, performances, diaries, student logs, etc.

(Refer to section 2.2.3, *Constructivism*)

- Formative assessment encourages the improvement of work in progress.

(Refer to section 2.3.3.3, *Formative assessment*)

- Opportunities for criterion-referenced assessment, in which students are evaluated on their own performance in relationship to clearly defined areas rather than in comparison to others, are provided.

(Refer to section 2.3.3.7, *Criterion-references assessment*)

- Assessment exercises are based on the goals of instruction and the material taught in class?

(Section 2.4, *Principles of assessment*)

The study of the assessment packages was beneficial for several reasons. It showed the alignment between standards-based curriculum and assessment practices. It provided clear evidence of the assessment methods that are being implemented by EFL teachers.

Since researchers, such as Grady (1998:25) believe that textual analysis is valuable when used together with other qualitative data-collection methods, the researcher chose to conduct focus group interviews in order to supplement the textual analysis.

### **5.7.1.3 Procedures used to analyse content**

In order to analyse the texts, the researcher considered several sets of guidelines. Smelser and Baltes (2001: 2697-8) offer the following suggestions:

- Data matrices are composed of rows and columns (Smelser & Baltes, 2001:

2697-8).

- A row is created for each sampled text (each assessment package is on a separate row).
- A column is created for each theme (the criteria that were used to analyse the assessment packages).
- The researcher examines the assessment packages to see if they comply with the criteria.
- Cells are filled in the matrix according to the occurrence of a theme and the rate of occurrence is recorded.

Mayring (2000) suggests the following steps for conducting qualitative textual analysis:

- Begin with a research question.
- Provide a theoretical based definition of the aspects of analysis. Establish the criteria for main categories and subcategories.
- Provide definitions, examples and coding rules for the categories based on theoretical suppositions. Collect the information in a coding schema.
- Revise the categories and coding schema during the analysis stage. Check the information for reliability during the process. A second coder was used to assure greater reliability.
- Conduct a final analysis of the texts. Check the reliability of the information at the final stage. A second coder made spot checks at the end of the analysis process.
- Interpret the results.

The researcher followed Mayring's model. The information was collected in a coding schema as follows:

- Each assessment package is assigned a code number that corresponds with the initials of the teacher who submitted the package.
- Each set of criteria is given a code number, beginning with the number 1. For example, number 1 – The tasks are specific and clear.

- The information in each assessment package is carefully analysed according to the criteria specified in section 2.6.1, *Criteria for effective alternative assessment*.
- The information from each assessment package is recorded in a computer file under the initials of the teacher.
- The information from each assessment package is organised into clear and specific categories that correspond with the code numbers 1 through 10. The creation of the specific categories is based on the material that has been submitted for analysis.
- The information from the 10 assessment packages is collated, compared, interpreted and analysed based on the predetermined criteria.
- The results of the context analysis are presented in clearly developed tables

### **5.7.2 Focus group interviews**

Powell and Single (1996:499) define a focus group as a group of individuals chosen and brought together by researchers to discuss and comment on the subject of the research using their personal experience. The benefit of focus groups is that they can produce a large amount of data in a short period of time (Cohen & Morrison, 2000:288).

Focus groups are an interview style intended for small groups (Berg, 1998:100). The focus group relies on interaction within the group based on subjects that are provided by the researcher (Morgan, 1997:12) and supplies insight into the attitudes, insights and opinions of the participants (Kreuger, 1994:19).

The focus group interview is a qualitative data gathering technique that finds the interviewer or moderator directing the interaction and inquiry in a very structured or unstructured way depending on the interview's purpose (Denzin & Lincoln, 1994:365).

### **5.7.2.1 Advantages of focus groups**

Focus group interviews are a form of semi-structured interviews that allow the interviewer to begin with an established set of questions, yet to digress from these questions in order to explore interesting or useful responses more extensively (Grady, 1998:20). Focus group interviews are a valuable tool for gaining insight into how teachers think and relate to specific subjects and ideas (Punch, 1998:174). An advantage is that they permit rich interaction and more personalised responses than is possible with questionnaires (McDonough & McDonough, 1997:184).

The group interaction in a focus group provides the participants with feelings of psychological security and spontaneity, which result in the stimulation and sharing of ideas (Stewart & Shamdasani, 1990:19). The interaction of the group stimulates participants to think beyond their own private thoughts and articulate attitudes and opinions (Kleiber, 2004: 91). Therefore, the interactive dynamics of a group can provide additional information that might not be disclosed in an individual interview (Grady, 1998:20).

As a result, there is a diverse range of responses that provide more extensive research (Merton, Fiske & Kendall, 1990:135). Not only do we learn what the participants tell us, we are also able to learn a great deal from how the respondents talk about the subject and how they interact with each other (Stewart & Shamdasani, 1990:15, 19; Grady, 1998:3).

An additional advantage is that focus groups help obtain general background information about a topic of interest as well as generating research hypotheses that can be submitted to further research and testing using more quantitative approaches (Stewart & Shamdasani, 1990:15).

In addition to recording verbal responses, researchers can take note of reactions during the interview. By later analysing the documents the researcher can glean greater insight into the interviewees' verbal responses by considering their body movement and facial expressions (Buhsmer, 2000).

The idea of conducting focus group interviews was very appealing to me. Since I am a teacher in the school system in Israel I am very interested in the practices and views of other teachers. I was anxious to discover the patterns of thought that would evolve from the discussions. I hoped that the data that was collected would contribute not only to my thesis but also to my teaching.

#### **5.7.2.2 Focus group participants and site of focus group interviews**

In order for focus groups to generate productive discussion it is imperative to choose an appropriate group of people to be included in the focus group (Anderson & Arsenault, 1998:201). This means choosing individuals with similar experiences or backgrounds who have the ability to interact with each other. The researcher selected both male and female teachers who would be able to give sufficient information and feel comfortable sharing their views and opinions.

Most focus groups consist of a small, limited number of people in order to encourage and enable all the members of the group to take an active part in discussions that should provide considerably more information than an interview with one person (Merton et al., 1990:137). There are differences in opinions regarding the ideal number of individuals for a focus group. Morgan (1998:10) and Krueger (2002:1) suggest six to eight participants as the ideal number, whereas Berg (1998:100) states that a focus group usually involves six to ten participants. Stewart & Shamdasani (1990:10) suggest eight to twelve participants as an ideal number.

The researcher held two focus group interviews. Anderson and Arsenault (1998:200) believe that two focus groups are sufficient in any study since additional groups rarely provide new information. The researcher decided to include six participants, which was a good number to facilitate group dynamics, yet not too many to prevent everyone from speaking. Since all the participants have had experience related to the subject, it was assumed that six people would be a sufficient number to carry on an in-depth discussion.

Consideration was taken to assure that the day of the week and the time of the interview would be convenient for everyone. This was done by asking the participants

to suggest dates and times that would be agreeable. The researcher then correlated the responses to find days and times for the group interviews. Both sessions were held in a conference room inside the library of a local high school, which was easily accessible to all the participants. The atmosphere was quiet and conducive to a group discussion. Light refreshments of cold drinks and pastries were available when the participants arrived. They were given the opportunity to meet and greet each other before the sessions began.

### **5.7.2.3 Questions for the focus group interviews**

The researcher asked ten open-ended questions in the focus group interviews. The questions were not given to the teachers beforehand. These questions were intended to encourage elaboration and discussion by the members of the group.

The questions provided a wealth of information that demonstrates how language teachers in Israel understand the principles of assessment based on a standards-based curriculum and the degree to which they have incorporated alternative assessment methods into their lessons.

The following questions were asked in the focus group interviews.

#### **Question 1**

How do you decide on the specific types of assessment activities that you use in your classrooms?

(Refer to sections 2.3.1, *Assessment and evaluation*; 2.3.2, *Testing and assessment*)

#### **Question 2**

In which way did the implementation of a standards-based curriculum influence your assessment practices in EFL?

(Refer to sections 3.2, *Definition of standards*; 3.5, *Kinds of standards*)

#### **Question 3**

To what extent do you still utilise traditional assessment methods and test items where knowledge is assessed – i.e. pen and paper tests and examinations which often include

multiple-choice, matching, true-false questions, paragraphs and essays?

(Refer to section 2.2.1, *Behaviorism*)

#### **Question 4**

To what extent does your assessment make provisions for the assessment of all the language skills?

(Refer to section 4.14.1.3, *Standards and benchmarks in assessment*)

#### **Question 5**

What classroom opportunities do your students have to be assessed by engaging in social interactions in authentic settings and situations?

(Refer to sections 2.2.4, *Multiple intelligence*; 3.5, *Kinds of standards*)

#### **Question 6**

What assessment methods and activities do you use on a regular basis?

(Refer to sections 2.2.4, *Multiple intelligence*; 3.5, *Kinds of standards*)

#### **Question 7**

How do you involve your students in classroom assessment?

(Sections 2.2.3, *Constructivism*; 2.5, *Traditional assessment*; 2.6, *Alternative assessment*)

#### **Question 8**

How do the assessment activities that you use encourage the improvement of work in progress?

(Refer to section 2.3.3.3, *Formative assessment*)

#### **Question 9**

What is your opinion on the use of criterion-referenced versus norm-referenced assessment?

(Refer to section 2.3.3.7, *Criterion-referenced assessment*)

#### **Question 10**

What are the main reasons why you might be reluctant to use alternative assessment

activities (such as projects, portfolios, student presentations and book tasks) in your classes?

(Refer to section 4.16, *The challenges of alternative assessment in EFL in Israel*)

The aim of the focus group interviews is not to reach an agreement, but to cover a range of opinions on the subject being studied (Kleiber, 2004: 100). The researcher's purpose is to gather extensive data in order to explore the research questions in depth.

## **5.8 FIELD NOTES**

Field notes refer to the researcher's detailed and descriptive documentation of the research experience that is gathered in the setting (the field) in which the qualitative study takes place (Anderson & Arsenault, 1998:126). Field notes include observations, a record of dialogue, personal reflection, a physical description of the setting, and decisions that are made that modify or guide the research procedure (Anderson & Arsenault, 1998:128). They are a major part of data collection in any qualitative study.

The field notes are written in a note pad during the discussion. They are reviewed immediately after the session and any missing details are added. Additional notes and thoughts are recorded after the meeting concludes so that they will not be forgotten. Anderson and Arsenault (1998:129) maintain that it takes three times longer to write good field notes than to do a real observation.

The field notes are transcribed into a computer file that has been set up prior to the first session. The file lists each question, which has been assigned a number; for example, the first question is number 1. The initials of the participants are used to identify their responses.

After all of the field notes are transcribed, the researcher carefully reads and considers the material. The responses are analysed with attention paid to the similarities and differences in the responses to each question. The researcher highlights comments that are particularly insightful and expressive to include in the final report.

### 5.8.1 Field analysis

Anderson and Arsenault (1998:132) offer several guidelines for field analysis:

1. Save time to analyse and reflect while in the field.
2. Begin to organise data and create initial coding schemes (e.g. setting, activities, perspectives, participation, relationships, and methods).
3. Make decisions in the field that force you to focus on your study.
4. Review field notes and listen to or transcribe cassettes.
5. Make a list of themes, patterns, and ideas or get intuitions that emerge.
6. Develop analytic questions and further define the problem.
7. Generate a list of questions or issues you want to clarify with respondents.
8. Make notes of personal reflections and decisions made, why they are made at this time, what is important.
9. Re-examine relevant theoretical literature.
10. Try to develop metaphors or analogies that help enhance your understanding.
11. Try to create a concept map that depicts preliminary constructs or relationships.
12. Remind yourself to keep looking and listening for what is *not* being said or witnessed.

Post-field analysis requires the analysis and interpretation of the raw data within the themes, patterns and categories that were recognised in the field analysis (Anderson & Arsenault, 1998:132).

## 5.9 DATA PROCESSING

Coding is very important for labeling and organising information. The researcher must consider several factors when coding and analysing data (Krueger, 1994:149-151). The following factors are essential:

- The researcher must consider not only the actual words but also the meaning of the words that the participants used.

- The researcher needs to consider the tone of the participant, how the tone changes and in what context it changes.
- The researcher should note any changes in opinions and attitudes that are brought about during the group interaction.
- The researcher should note which areas of the discussion elicit the strongest responses and note the significance.
- The researcher should pay attention to more specific rather than vague responses.
- The researcher should not concentrate too closely on analysing every individual at the expense of bigger ideas.

Data analysis involves more than providing an explanation of the data that is collected. Qualitative data needs to be organised, so that related information can be selected and separated from information that is not directly connected to the study (Wiersma, 1995:216). In order to produce the most effective results, the researcher must have the necessary skills in order to be able to condense, reorganise and collate related information from the data collected in the study.

The field notes of each session were hand-written and transcribed immediately after each meeting. Measures were taken to ensure that all the data collected at the meetings was reviewed and transcribed. The transcripts of the discussion went through careful analysis and interpretation.

In addition, all focus group interviews were audio-taped (with the permission of the participants) as the researcher agrees with Patton (1990:348) who believes that a tape recorder is an essential tool for assuring that all the conversations during the interview are on record.

The decision to tape each session was made for several reasons:

- The researcher felt that it was necessary to listen to the responses of the participants several times in order to interpret and analyse them effectively.
- It was felt that additional attention to the verbal responses of the group would

yield a better understanding and richer conclusions.

- There was a concern that strict concentration on taking notes might detract from the researcher's observations of the group.

The tapes were listened to several times in order to transcribe them accurately. The transcripts were reread numerous times so that important information would not be excluded. Transcriptions of the focus group interviews are included as appendixes (Appendix 4; Appendix 5).

A filing system was set up to manage the raw data. Back-up files and copies were made of vital information to prevent any loss of information.

Data was organised into smaller units based on ideas, themes, categories and subcategories in order to shed light on the findings of the research. A “master list” for coding was devised and is included as an appendix (Appendix 6).

### **5.9.1 Segmenting**

Segmenting entails dividing data into significant analytical components (Johnson & Christensen, 2001:502). These units are comprised of words, sentences or passages that convey a specific meaning that might be important for the research study. The transcripts were segmented to facilitate analysis and interpretation.

### **5.9.2 Coding**

The coding process involves marking segments of text data with symbols, descriptive words or category designations (Johnson & Christensen, 2001:503). Segmenting and coding interact with each other. Coding isolated central ideas and themes were essential for analysis and study.

### **5.9.3 Assembling a master list**

A list of the codes used in the research study should contain each code followed by

the full code name and a concise definition or description of the code (Johnson & Christensen, 2001:504). This master list should be easy and clear for other researchers to use (Appendix 6).

In this chapter the data generated from the focus group interviews with a total of twelve EFL teachers and the analysis of the assessment packages of ten Grade 11 EFL teachers are presented.

#### **5.9.4 Enumeration**

Enumeration is the practice of quantifying the data that has been collected. This provides information regarding the "frequency" of specific data that can be useful in the research report (Johnson and Christensen, 2001:505).

#### **5.9.5 Intracoder reliability**

In order to check for intracoder reliability, the researcher asked a colleague to analyse some of the assessment packages. The researcher then compared her results with the results of the second analysis. The decision was made to employ an additional coder since reliability data results from duplicating the coding, categorising and analysing a sample of data that is collected (Krippendorff, 2004b:414).

#### **5.9.6 Broad categories**

The data from the textual analysis yielded the following six broad categories:

- Bagrut modules and assessment
- Traditional assessment of recall and rote learning
- Rubrics and writing
- Projects and book tasks
- Assessment of integrated skills
- Traditional assessment test items dominate assessment

The data from the focus group interviews yielded the following seven broad

categories:

- Bagrut modules and assessment
- Traditional assessment of recall and rote learning
- Rubrics and writing
- Time, class size and assessment
- Projects
- Assessment of integrated skills
- Traditional assessment

## **5.10 SUMMARY**

This chapter discussed the empirical stage of the study. Attention was paid to research design, research methodology, sampling, data collection and data processing. The qualitative research methods that are employed in the study - textual analysis, focus group interviews are identified and explained. The research findings are presented in the next chapter that will also include conclusions, implications and suggestions.

## CHAPTER 6

### RESEARCH FINDINGS AND CONCLUSIONS

#### 6.1 INTRODUCTION

In this chapter the data generated from the focus group interviews with a total of twelve EFL teachers and the analysis of the assessment packages of ten Grade 11 EFL teachers are presented.

Firstly, the procedures for conducting the textual analysis and the focus group interviews are briefly discussed. The researcher had assured participating schools as well as the participants of confidentiality and anonymity. For this reason the names of the participants or the schools where the interviews were held are not included in the discussion.

The empirical study focused on the research questions and a set of related questions.

The research question is:

How do EFL secondary school teachers in Israel implement standards and the principles of alternative assessment methods (that were discussed in the literature study chapters 2-4 of this thesis) in their assessment practices?

Related questions are

1. How do assessment activities and methods reflect the theories and principles that underlie assessment practices?
2. To what extent did EFL teachers' awareness of standards and alternative assessment influence their assessment practices?
3. What is the implication of alternative assessment in EFL teaching in Israel?

The aim of the empirical study was to identify the activities and methods that English

teachers in Israeli high schools employ to assess their students and to determine what impact the implementation of standards had on assessment in EFL.

The study also investigated teachers' perceptions of assessment and alternative assessment in teaching EFL in Israel.

## **6.2 PROCEDURES FOR CONDUCTING THE TEXTUAL ANALYSIS**

A textual analysis was carried out of the assessment packages of ten English teachers in EFL who are currently teaching. They included samples of the tests and assessment methods that they used in their classrooms to assess their students over a semester. Each teacher also indicated the percentage breakdown of the students' final semester grade.

The researcher organised the data in the assessment package according to the criteria for effective assessment that was previously discussed in this paper (section 5.7.1.2, *Criteria used to determine whether alternative assessment measures were incorporated*).

The assessment package of each teacher was analysed separately and the researcher recorded the data on a chart. The data was then reorganised and related information was selected while unrelated data was discarded. It was necessary to reduce the amount of data in order to arrive at an accurate description and interpretation of the research subject (Wiersma, 1995:216).

After the ten assessment packages were studied, the researcher compared the data from all of the assessment packages and designed a chart that displayed the combined results of the textual analysis (Appendix 1; Appendix 2).

The ultimate goal of the textual analysis was to show how teachers incorporate alternative assessment into their classroom teaching. The researcher was able to answer the research question by reviewing, comparing and analysing the contents of the assessment packages.

### **6.3 PROCEDURES FOR CONDUCTING FOCUS GROUP INTERVIEWS**

The first focus group session was held on Monday afternoon, June 5, 2006 from 5:00 p.m. until 6:30 p.m. The second session took place on Sunday afternoon, June 11, 2006 from 5:00 to 7:00 p.m.

Four of the six participants in the first group session had previously met each other and three of the participants in the second session were colleagues. The short time spent at the refreshment table allowed the teachers to speak informally.

The chairs were arranged in a circle so that all the participants would be able to see each other and have equal access to the discussion. An audio-tape was in full view and the reason for taping the interview was explained.

At the beginning of each session, the interviewer welcomed the teachers and thanked them for their time and assistance, and ensured them that the content of the interviews would remain confidential. The researcher initiated each session by introducing the objectives of the session and explaining the protocol in order to put the participants at ease.

An audio-tape recorded the full interviews in order to help the researcher later review the texts that had been taken during the interview. The audio-tapes enabled the researcher to concentrate on vocal intonation which intensified the meaning of the interviewees' statements. During the sessions the researcher quickly jotted down notes related to facial expressions and gestures that might be useful in the final analysis. Field notes were written in a note pad during and directly after each session. These notes consisted of comments that were made during the interviews as well as observations that the researcher made. The interviewer also recorded her impressions after the interviews in order to remember particular details that were said and other thoughts that might be significant later. Immediately after each session, the recorded texts were transcribed to facilitate segmentation and coding. Complete transcripts were made of the interviews and follow-up comments of the participants (Appendix 4; Appendix 5). A master list was compiled of the important code words and their implication (Appendix 6).

The researcher proceeded by asking the questions that were based on specified criteria listed in section 5.7.2.3, *Questions for the focus group interviews*. The researcher began with simple questions and progressed to more complex questions after the participants appeared comfortable and willing to participate in the group discussion as suggested by Stewart and Shamdasani (1990:61). The researcher was careful to ask open-ended questions, including words like how, why, in order to allow for greater discussion (Stewart & Shamdasani, 1990:65).

The teachers seemed a bit reluctant to offer information at the beginning of each session; however, they did speak more as the sessions progressed. The participants of the second focus group interview appeared more passionate about the subject. As the time went by, a few of the teachers seemed anxious to express some of the frustration that they felt regarding assessment. They seemed pleased to have an opportunity to find a release for some of their dissatisfaction with the current teaching system and assessment guidelines. This is particularly apparent in the transcript of the second focus group interview (Appendix 5).

The style of interaction during the interviews was informal, utilising open-ended questions, with many opportunities for examples and elaborations. The meeting was constructive; participants were asked to be open; and all members of the group had the opportunity to contribute to the discussion. Questions were repeated or explained whenever any misunderstanding or uncertainty was detected. Within the one to two hour interviews, the teachers exhibited a willingness to answer the questions and share their views, although some were more actively involved than others and more eager to express their ideas. Teachers presented different views during the discussion although they were in agreement on most issues.

During the interviews the researcher paid attention to the non-verbal communication of the participants, such as gestures, facial expressions and body language. The notes that were carefully written during the length of the interviews served as a valuable addition to the recorded comments of the participants. Significant patterns emerging from the analysis of the participants' responses were synthesised and related to prior research.

## **6.4 PRESENTATION OF THE FINDINGS OF THE TEXTUAL ANALYSIS**

The textual analysis provided an insight into the assessment practices of a sampling of EFL high school teachers. The examples of assessment evident in each of the packages are interrelated to the percentage distribution of the semester grade (Appendix 3). This is obvious in the following description of each assessment package.

### **6.4.1 The composition of the assessment packages**

- **Assessment package 1**

Thirty percent of the first assessment package is based on tests. These tests are comprised primarily of traditional assessment test items (multiple-choice; closed questions). There is a short composition on each test that follows the strict Bagrut model. These tests are identical to the Bagrut modules that the students are taking at the end of the semester. Seventy percent of the semester's activities are alternative assessment activities, such as literature analysis, project work and oral presentation of the project. These activities integrate language skills and require students to engage in higher order thinking. These activities assess what students know and are able to do with their knowledge and skills.

- **Assessment package 2**

Fifty percent of the second assessment package is awarded to tests which are modeled on the Bagrut module that the students are taking. These tests are comprised solely of traditional test items, such as multiple-choice, true/false and short fill-in answers. An additional ten percent is given for classwork and homework that are primarily traditional assessment exercises. Forty percent is divided between literature assignments and a project. These alternative assessment activities allow students to be creative and work independently or to work with peers

- **Assessment package 3**

Forty percent of the third assessment package is allotted to tests which are identical to the Bagrut module that the students are taking. These tests consist of traditional assessment test items. Ten percent is given for listening comprehension activities that are multiples choice or fill-in. An additional ten percent is awarded for homework and classwork which are traditional assessment type exercises. Forty percent is divided between a project and book tasks. These activities provide students with the opportunity to choose topics and use authentic language.

- **Assessment package 4**

Sixty percent of the fourth assessment package is allocated for traditional assessment type tests that mirror the Bagrut module. An additional ten percent is comprised of listening activities that are based on traditional assessment test items, such as multiple-choice and fill-in questions. Fifteen percent is awarded for classwork and homework that is primarily traditional assessment type exercises. Fifteen percent is given for short comparative literature assignments which provide the only evidence of alternative assessment.

- **Assessment package 5**

Thirty percent of the fifth assessment package is assigned to traditional type tests similar to the Bagrut module. Fifteen percent is given for classwork and homework which is comprised of traditional assessment type exercises. Fifty-five percent is divided among literature assignments, a project, an oral presentation and book tasks. These activities provide opportunity for alternative assessment and allow students to engage in higher order thinking tasks.

- **Assessment package 6**

Forty percent of the sixth assessment package is comprised of traditional assessment test items such as, multiple-choice, binary questions and fill-ins. Ten percent is given

for listening comprehension practice that is comprised almost completely of traditional assessment question items, such as multiple-choice and fill-ins. Fifty percent of the semester grade is divided among literature assignments, a project and an oral presentation.

- **Assessment package 7**

Thirty percent of the seventh assessment package is designated for tests. These tests are comprised primarily of traditional assessment test items (multiple-choice; closed questions). These tests are identical to the Bagrut modules that the students are taking at the end of the semester. There is a short composition assignment on each test that follows the strict Bagrut model. Fifty percent is given for classwork and homework that is comprised essentially of traditional assessment type activities. Twenty percent of the semester's grade is assessed by alternative assessment (book tasks). Book tasks increase the students' learning potential by challenging them to employ their knowledge and skills to read and analyse a piece of literature, compare genres and produce an original piece of written or oral work.

- **Assessment package 8**

The eighth assessment package relies essentially on traditional assessment. A high seventy percent of the semester grade is allotted for tests (multiple-choice, binary and short fill-in questions) that mirror the Bagrut module. Thirty percent of the semester grade is given for book tasks which offer students the opportunity to be assessed by alternative methods. These activities encourage creativity and allow students to construct their own meaning from books they have read.

- **Assessment package 9**

Thirty percent of the ninth assessment package is assigned to tests. These tests are comprised primarily of traditional assessment test items (multiple-choice; closed questioned). There is a short composition on each test that follows the strict Bagrut model. Twenty percent of the semester grade is made up of classwork and homework

which are also primarily traditional assessment type activities. Fifty percent of the grade is given for project work and an oral presentation of the project. These activities allow for alternative assessment which encourages the use of authentic language and encourages higher order thinking.

- **Assessment package 10**

Sixty percent of the tenth assessment package is awarded to tests that are modeled on the Bagrut modules. An additional ten percent is given for classwork and homework that are similarly designed. Thirty percent is given for book tasks and an oral presentation of the books tasks.

There are several generalities that can be drawn from the textual analysis. Although the percentage of the grade that was given for tests varied in each assessment package, all the tests were based exclusively on the model of the Bagrut module. The assessment packages clearly indicated that the Bagrut module that the students were taking at the conclusion of the semester strongly influenced the teachers' assessment choices. The Bagrut modules do not in general facilitate alternative assessment. Only certain modules include a written presentation (a composition) that requires students to analyse a situation and construct a written piece of work. This written presentation comprises thirty to forty percent of those modules; sixty to seventy percent is based on traditional assessment test items. Only a few of the assessment packages included assessment for the modules that contained writing. Most of the assessment packages that were analysed included almost one hundred percent traditional type of questions (reading comprehension and listening comprehension). Yet, a few of the modules require written presentations in addition to the reading comprehension exercise that is mainly based on traditional assessment. The written composition does challenge students to analyse a situation and design a response. However, the written presentation is part of only three modules and there were very few examples of written composition in the assessment packages. Examples of these compositions are:

**Written Presentation:** The municipality has announced its plans to build a large shopping mall near your home. Write a formal letter to the municipality expressing your support of or objection to the plan. Give reasons for your position.

The written presentation provides a limited opportunity for students to relate to a real life situation and create a piece of original work. Unfortunately, the number of words is limited to between 120-140 words which limit the students' opportunity to express themselves freely and reduce writing to a technical exercise. The modules, therefore, have a negative effect on teaching and assessment.

#### **6.4.2 Analysis of the assessment packages**

In this section will be discussed the assessment packages according to the criteria that were identified in section 5.7.1.1, *Criteria used to examine and judge the influence of standards on assessment*.

##### **6.4.2.1 The influence of standards on assessment**

- The extent to which assessment activities demonstrate what the students are required to know and be able to do

Not all of the assessment packages showed evidence of the influence of standards on assessment. The assessment packages of four teachers did not expect learners to demonstrate what they know and can do. Assessment in these four packages consisted mostly of traditional assessment activities in which learners were not expected to apply knowledge and skills, but instead were merely required to recall information. They include multiple-choice questions that students can answer correctly by guessing at the right answer. The short fill-in questions ask for recall from the reading passage and do not challenge students to construct their own answers or build on the information they have read.

Project work is, however, evident in the assessment packages of six teachers. These assessment tasks require students to demonstrate by means of practical presentations and group discussions what they know and are able to do with their knowledge and skills. Unfortunately, there is only one project in each of these assessment packages. Therefore, the opportunity for students to use combined knowledge and skills is limited. As a result, there is limited occasion for their competence to be assessed in an authentic task. Examples of projects are:

**Project work 1:** Greater and Lesser Authors: Choose a well-known playwright or poet (not an Israeli person). Use the Internet to research the personality that you have chosen and be prepared to give background information on the writer to the class. You may work in pairs or threes but each of you should have a part in the class presentation itself.

**Project work 2:** Discrimination and Prejudice Around the World: Cut out or download a relevant newspaper or Web article. Read the article and bring it to class. Be prepared to discuss the article with the class.

Book tasks were included in only half of the assessment packages and literature assignments were evident in five of the assessment packages. These tasks expected learners to transfer themselves to the situations, feelings, thoughts or actions of the characters in the text, rather than answering numerous short answer questions on the text. In many cases learners were expected to express their own opinions and come to conclusions or make suggestions as is evident in the following example, where learners had to suggest a plan of treatment for a specific character.

**Literature assignment 1:** Read James Joyce's story *Eveline*. Imagine that you are the psychiatrist treating Eveline. Write an essay in which you describe her condition (based on what you know about her behavior) and suggest a plan of treatment

**Literature assignment 2:** Read *My Papa's Waltz* by Theodore Roethke and *Walking Away* by C. Day Lewis. Write an essay explaining what you think characterises a good relationship between a parent and a child. Use reference to both poems in your essay.

Book tasks, literature assignments and project work expect students to integrate knowledge and skills and demonstrate by means of oral presentations, essays and group discussions what they know and can do. The inclusion of these aspects is, therefore, an indication of the influence that standards had on assessment practices. However, activities that challenge students to use their knowledge and skills in authentic situations are a small part of the assessment evident in the assessment packages. In those assessment packages that included alternative assessment activities

there was only one project, one literature assignment and one book task. The bulk of assessment was comprised of traditional assessment exercises.

- The extent to which tasks are an exemplary demonstration worthy of emulation

Only a few of the assessment packages had tasks that are worthy of emulation. Tasks that are worthy of emulation are tasks that require students to simulate real-life situations with real people in authentic settings. These tasks require students to act in situations similar to situations that occur in real life. For example:

### **Book Task 1:**

You are talking to a friend who plans a trip overseas. While listening to him, you realise that the book you have just finished reading can give your friend some excellent background about the places he intends to visit, the people he is going to meet, the customs etc. Write down such examples that are described in the book and explain why it would be wise for a tourist who is going to that place to read the book before setting out on the trip.

### **Book Task 2**

You are trying to convince the librarian at your school or in your neighborhood that buying the book for the library is really worthwhile. However, the librarian has reasons of his or her own to refuse your request. Write out the conversation between the two of you bringing up relevant points.

These tasks are worthy of emulation because there is a possibility that learners will come across similar tasks in real life. Therefore, these activities can provide them with a practical experience that might be useful in their adult life. There were a few examples of tasks of this nature in the assessment packages.

- The extent to which tasks are evidence of the vital basic knowledge rather than rote learning

Only six of the packages showed tasks that called for basic knowledge rather than rote learning. There were only one to two examples of tasks like these in those packages. However, there was a lot of evidence of rote learning found. This was apparent in eight assessment package that showed evidence of from fifty percent to eighty percent rote learning.

There are a limited number of examples of tasks that require the use of basic knowledge rather than rote learning. For example:

**Project 1:** Choose a famous (or less-known) personality who has made a major contribution to society. Write ten questions that you would like to know (or questions that you think your classmates would like to know) about that person. Using at least two search engines, surf the Internet and find the answers to your questions. Pretend that you are that person and write six to ten diary entries in which you incorporate the information that you find.

**Project 2:** Building Walls –After reading the poem *Mending Wall* by Robert Frost choose a famous wall that either existed (Berlin Wall, Great Wall of China) or that exists today (USA/Mexico border, separation fence in Israel). Use the Internet to research your topic and write a report. Refer to the poem in your essay. Relate to the following questions:

1. Where is the wall?
2. Why was it built?
3. What is the wall made of?
4. What do people think about the wall?

These tasks require learners to use their basic knowledge to research a topic, use the Internet and collect, analyse and compare information. They need to use research skills, to paraphrase and organise their thinking. They require basic knowledge and skills for communication. These tasks challenge learners to use the knowledge that they have to complete complex tasks,

- The extent to which tasks encourages cognitive growth.

Standards describe the level at which learners should demonstrate their achievement. There should be conceptual progress and assessment activities should provide for and encourage cognitive growth. Only a limited number of the assessment packages included activities in which cognitive growth are encouraged. Project work in particular is used for this purpose.

There are examples of project work in the assessment packages that encourage cognitive growth. The following are examples of this:

**Project 1:** Class project: Natural Disasters. Choose any natural disaster and choose three of the following four components to include in your project:

1. Write a research report on your subject.
2. Write a speech explaining why the disaster that you chose is the most devastating.
3. Prepare a short brochure to give information about the disaster (use graphics and text).
4. Write four diary entries written by someone who survived the disaster.

**Project 2:** Satire: Use the Internet to research various types of satire. Choose one type of satire. Prepare a short paper in which you include an explanation and examples of the particular type of satire that you chose. Explain why it is popular.

Projects present learners with the challenge of facing a new subject and using their knowledge and skills to investigate, interpret and produce something new. Only six of the assessment packages included project work.

- The extent to which tasks maximise students' learning potential

There are a few examples of assessment tasks that maximise the students' learning potential. For example:

**Task 1:** After reading the play *Visit to a Small Planet* find a web site that you think has a connection to the play. Write an essay between 100-200 words in which you discuss three to four important points in the play. Make sure to use at least three quotes from the play to prove your points. Add the web link to your essay.

**Task 2:** Israel and Identity. Read the excerpt *My Enemy Myself* by Yoram Binur. Find an article or published incident that you think is relevant to that discussion. Print it out and bring it to class. Be prepared to give a three to five minute speech to the class about the article that you have chosen.

These tasks maximise the students' learning potential by challenging them to use their knowledge and skills to read and analyse a piece of literature, compare genres and produce an original piece of written or oral work. This type of task, however, was evident in only half of the assessment packages.

- The extent to which tasks are specific and clear and can be easily assessed

There were clear and specific instructions for projects in six assessment packages. These projects were assessed with the rubric suggested in the *NBA Handbook* (Ministry of Education, Culture and Sport, 2003:39). The rubric for assessing written presentations is accompanied by clear descriptors and is comprised of the following categories:

1. Content and organisation
2. Vocabulary
3. Language use
4. Mechanics

Five assessment packages included assessment for oral presentations of the projects. These presentations were assessed by the project presentation and interaction rubric suggested in the *NBA Handbook* (Ministry of Education, Culture and Sport, 2003:36). The categories assessed in the presentation are:

1. Project presentation
2. Accuracy
3. Project interaction

Rubrics are usually used when performance-based assessment is done. Learners in most cases were given the rubrics beforehand. They knew the criteria for assessing their work prior to beginning their tasks. This facilitated proficiency since it was clear what criteria needed to be met in order to reach the standards.

- The extent to which tasks require students to use their knowledge and skills in authentic situations

Authentic assessment means that teachers need to give learners tasks and problems that they are likely to encounter in real life. This is best done by placing learners in simulated situations and then requiring them to complete certain tasks. This type of assessment tasks was found in a large number of the assessment packages. However, each assessment package had a very limited number of these assignments, usually only one of a particular type.

Many of the book tasks provide assessment tasks that require students to use their knowledge and skills in authentic situations. For example:

**Book Task 1:** The book you have just read had such an impact on you that you have decided to join a group of activists fighting for or against the issue referred to in the book. Explain the issue, explain why it is important enough to fight for or against it and suggest at least two different ways of doing it.

**Book Task 2:** Imagine that one of the characters in the book was brought to trial for a certain crime or misconduct. Explain why you think it is justified to bring the person to court and what you believe his or her verdict according to the Israeli law should be.

The inclusion of authentic assessment in a large number of assessment packages gives an indication of the type of influence that standards had on assessment practices. However, the fact that a relative small number of these tasks are found in the packages indicates that they are not the main assessment method.

- The extent to which task requirements are flexible, allowing for student diversity

The book tasks and literature assignments allowed students to select from a variety of activities. This was evident in most of the assessment packages that had book tasks and a small number of the assessment packages that contained literature assignments. Learners were given a number of diverse activities that they could choose from. For example:

**Task 1:** Extensive Reading Activity. Complete one of the following:

- Draw a summary of the story in comic strip form.
- Turn your book into a picture book for young children with short sentences under each picture.
- Produce a creative time line of the events in the story on a poster.
- Create an advertisement for your book for either TV (a video) or the radio (a tape)
- Turn your book into a mind map

**Task 2:** Literature Assignment: Choose one poem that was studied in class this semester. Present either a written or an oral presentation on the poem. Discuss any four of the following features of the poem:

1. Structure
2. Language
3. Rhyme scheme
4. Background of the poet
5. Relevance to teenagers
6. Background of the poet

The choice of activities enabled learners to match their individual interest and learning style to the assignment. However, only a limited number of opportunities were available for this type of assessment in the assessment packages.

- The extent to which tasks allow for students to utilise their multiple intelligences

Only half of the assessment packages contained book tasks, yet all of these book tasks allowed students to use their multiple intelligences and choose from a variety of assessment tasks. This was evident in the choice of book tasks that were given. Examples of book tasks:

**Book Task 1:** Design a wardrobe for the main characters in your book.

**Book Task 2:** Find music or write your own music that suits the background of the story. Present a tape or disk with the music.

**Book Task 3:** Choose an important scene in the book. Act out the scene.

**Book Task 4:** Write a letter to one of the characters expressing your opinion about his or her conduct in a certain situation.

The choice of book tasks allowed learners to call upon their multiple intelligences and choose a task in an area where they felt confident of achievement. However, there was evidence of book tasks in only half of the assessment packages and there was only one book task in a semester.

#### **6.4.2.2 The incorporation of alternative assessment methods**

It was apparent from the assessment packages that all of the teachers utilised traditional assessment. The weight of traditional assessment ranged from thirty percent to as high as eighty percent of the semester grade. This is evident from Appendix 3.

- The use of traditional assessment

An investigation of the assessment packages reveal that a considerable percentage of the semester grade is based on tests (Appendix 3). In every assessment package the tests were composed almost essentially of traditional assessment test items, such as fill-in short answers and multiple-choice items. Examples of these types of questions are:

1. What does Lu explain in his second answer (line 16-21)

Complete the sentence

He explains.....

2. What information are we given in lines 1-7?
  - i) The reason that Lu is good at his job
  - ii) The reason that Lu is being interviewed
  - iii) The history of Wonton Foods
  - iv) The history of fortune cookies

Traditional assessment relies heavily on recall and recognition. It tests memory and requires learners to recall information that they have already seen. It doesn't challenge students to use their knowledge and skills to construct new meaning. It encourages learners to merely recognise answers. Traditional assessment doesn't assess true understanding nor train learners to transfer their knowledge and skills to diverse situations.

- Encouragement of memorisation and rote learning of isolated bits of information

Three of the assessment packages show weekly vocabulary and spelling tests and monthly grammar tests. These tests require memorisation and rote learning of grammar rules out of context and translation from English to Hebrew. For example:

1. Translate the following words into Hebrew:

- a) border
- b) permission
- c) property

2. Choose the most suitable word or phrase to complete the sentence:

- 1) Why ..... today? You don't look well.
  - a) you are working
  - b) you work
  - c) are you working
  - d) you worked

Traditional assessment, primarily of memorisation and rote learning was evident in all the assessment packages. It was the main assessment in every package, whereas there were very few opportunities for alternative assessment.

- Students are assessed on oral activities

The only evidence of assessment of oral activities was in five assessment packages.

Students gave oral presentations of their projects and were assessed according to rubrics suggested in the *NBA Handbook*. The rubric for assessing the project presentation interaction included criteria in the areas of: project presentation, accuracy and project interaction. Learners were required to present their projects using accurate language and to answer questions about the project.

- Students are actively engaged in assessment

In only three cases was there evidence that students were actively involved in assessment. This was apparent in one book task where students were required to complete a check list prior to submitting their assignment and in two projects where students were required to write their reflections.

In a limited number of cases students were made aware of the criteria for assessment before beginning their tasks. Six assessment packages contained projects that were assessed with the rubric suggested in the *NBA Handbook* (Ministry of Education, Culture and Sport, 2003:39). The rubric for assessing written presentation is accompanied by clear descriptors and is comprised of the following categories:

- Content and organisation
- Vocabulary
- Language use
- Mechanics

Five of the assessment packages contained assessment for oral presentation of the projects. The oral presentations were assessed by the project presentation and interaction rubric that is suggested in the *NBA Handbook* (Ministry of Education, Culture and Sport, 2003:36). The categories assessed in the presentation are: project presentation, accuracy and project interaction. Students were shown the rubrics beforehand so they were aware of the criteria they needed to meet in order to succeed in the task.

- Opportunities for self-assessment and peer assessment

Opportunities for self-assessment was limited to only one book task where students were required to complete a check list prior to submitting their assignment and in two projects where students were required to write their reflections. Reflections enable learners to review, consider and evaluate their learning. Reflections allow them to think about how they might do something differently in the future.

- A provision is made to assess all language skills

The assessment packages revealed limited provisions to assess all language skills. Only three assessment packages showed assessment of listening comprehension. There was evidence of the assessment of spoken skills in only half of the packages. Reading comprehension was assessed, albeit according to traditional assessment, in all the assessment packages. A combination of writing in project work, literature and book tasks were seen in most of the assessment packages. However, there was generally very few of these writing tasks in a semester.

- Encouragement of varied mediums for assessment, such as: journals, projects, performances, diaries, student logs, etc.

There was evidence of project work in six of the assessment packages. Five assessment packages contained oral presentations. Five assessment packages had book tasks and five packages contained literature assignments. Although it is applaudable that alternative assessment activities were included in some classroom assessment, this is insufficient. There was only one project, one book task and one literature assignment in the packages that included these activities. Alternate assessment should be the predominant assessment according to the new curriculum; however, traditional assessment is prevalent in almost all the assessment packages that were analysed.

- Opportunities for formative assessment

Formative assessment was evident only in project work and book tasks. Students submitted first drafts of these assignments that were then reviewed by their teachers. They were then given opportunities to review the corrections and suggestions and resubmit a final draft for grading. Their work was thus assessed while in progress and then again at completion. Formative assessment promotes assessment for learning.

- Opportunities for criterion-referenced assessment

Students had the opportunity for criterion-referenced assessment only in project work, oral presentations and in one or two writing assignments. Evidence of this was found in only six of the assessment packages and only in half of the packages for oral presentations. In these cases there was only one project and one oral presentation. Therefore, students were assessed by criterion-based assessment only once or twice a semester.

A rubric was used to assess projects, oral presentations and writing assignments. Students were shown rubrics beforehand for both written presentation and project presentation. Half of the teachers showed an understanding of the concept of criterion-based assessment. This is evident in the examples of rubrics that were found in the assessment packages. The rubrics in the assessment packages were the same rubrics suggested in the *NBA Handbook* for assessing project work and written presentations. These rubrics clearly specified the skills and abilities that were the criterion for assessment. Criterion-referenced assessment was rare, whereas the assessment packages indicate that traditional assessment was the constant and continuous form of assessment.

- The use of assessment exercises that are based on the goals of instruction and the materials taught in class

The goal of instruction evident in the assessment packages was to prepare students for the Bagrut module that they were taking that semester. The tests were modeled on

that module. Therefore, the tests reflected the type of Bagrut questions which are primarily traditional type of questions. Although there is a link between the goals in the curriculum and the assessment tasks, most of the assessment evident in the assessment packages was not related to the assessment goals of the curriculum. For example, assessment should be formative and summative, yet most of the assessment was based on the results of tests and exercises that were comprised of traditional assessment test items. The standards in the curriculum call for a diversity of assessment methods including assessment tasks in which students can display their knowledge. These types of assessment tasks occurred infrequently in the assessment packages. The curriculum stresses development assessment showing progress towards reaching the standards and assessment tasks which are included within the learning-teaching-assessment process. However, almost all of the assessment in the packages pointed to assessment essentially of a final product. This was evident in the tests and classroom and homework exercises.

The literature assignments and book tasks were intended to develop critical thinking. One of the assessment goals of EFL is to develop critical thinking. Alternative assessment tasks place greater emphasis on problem-solving, higher level thinking and reasoning. These tasks encourage students to use their skills and knowledge to interpret and analyse.

There are several examples of these tasks:

**Book Task 1:** Poems that Talk – Search your literature book or other poetry books on the Internet for a poem (not a song) that really 'speaks' to you. Print it out (with the poet's name) and attach a 150 word essay (Bagrut standard) explaining what you think the poem is about and why, and the reason you like it. Be prepared to read and discuss your poem in class.

**Book Task 2:** As the author of the book you have been asked to write a continuation to the book – a second volume. Which aspects of the book would you happily develop in the next volume? Which aspects would you choose to ignore? Explain.

These tasks encourage students to use their cognitive and language skills to stimulate creative thought and arrive at conclusions.

## **6.5 PRESENTATION OF THE FINDINGS OF THE FOCUS GROUP INTERVIEWS**

The two focus group interviews provided an insight into EFL teachers' beliefs about assessment.

### **6.5.1 Analysis of the focus group interviews**

The focus group interviews yielded considerable information about the understanding and attitude that teachers have regarding assessment. The report of this information is presented in the following sections.

#### **6.5.1.1 How decisions on the specific type of assessment activities to be used are made**

Although the teachers were asked how they decided on the specific assessment they used, the question immediately generated a discussion about the role of assessment in learning. Some of the teachers stated they did not consider assessment to be very important. The reasons that were given for this was that assessment was negative in half of all situations and that the focus should, therefore, not fall on assessment. One of them stated: "In my classes assessment is pushed into the background". He said that he believed that assessment was used essentially for the purposes of parent-teacher conferences and assigning report card grades. However, he did acknowledge the necessity to assess "big things" like projects and book reports.

Several of the teachers maintained that the main objective of assessment was to check how pupils used what they had learned. According to one of them assessment "helps us see just how much they {students} assimilate what we put into them". Only one of the teachers disagreed and insisted that assessment should not be used to see what a student knows, rather assessment should be used when students are doing well in order to build their self-confidence, to offer encouragement and "to boost their egos".

From the responses of a number of teachers it appeared that the specific type of assessment was inter alia influenced by the skills they wanted to assess. Most of them, however, admitted that the skills that they taught and assessed were determined by the

Bagrut modules that their students were taking. For example, one of the teachers explained that she used alternative assessment with 9<sup>th</sup> and 10<sup>th</sup> graders since they were more willing to be creative and imaginative. However, she relied on traditional assessment with 11<sup>th</sup> and 12<sup>th</sup> graders because that was the type of assessment in the Bagrut.

From the responses of a number of teachers it appeared that the specific type of assessment was also influenced by the level of the students' comprehension and ability. The age of students also affected the choice of assessment. Two more teachers stated that the type of assessment they chose was determined by the skills that they wanted to assess and that was influenced by the Bagrut modules. Another one added that the choice of assessment was also determined by the different types of students. For example, some students needed to be assessed orally because writing was difficult for them.

Time considerations also affected the choice of assessment. Traditional assessment activities were used primarily because they took less time to prepare and were easy to execute.

## **Conclusion**

*Ignorance regarding the importance of assessment and the link between assessment and learning:* It is apparent from the teachers' responses that they do not perceive the connection between assessment and learning. As a result, assessment does not play a prominent role in instruction.

*The role of the Bagrut modules:* Many teachers said that the assessment they chose was determined by the skills that they wanted to teach. Unfortunately, those skills were dictated by the format of the Bagrut modules. Consequently, "teaching to the test" exerted a powerful influence on classroom assessment. This thesis has discussed the detrimental effect that such a practice has on learning. (section 3.6.3, *Teaching to the test*).

The majority of the assessment packages similarly revealed that teachers used assessment to prepare their students for the Bagrut modules. This was evident in all the tests that mirrored the traditional test items that comprise Bagrut modules, such as multiple-choice, fill-in and short answer questions.

*Time considerations:* Traditional rather than alternative assessments is used mainly because they are easy to administer and time-saving. The assessment packages reveal many traces of traditional assessment. It is in fact the principal means of classroom assessment. These assessment packages confirm what the teachers said in the focus group interviews.

*Learners' level of development and abilities:* It seems that the learners themselves are taken into consideration when assessment activities are decided on. Teachers make provision for age differences, level of comprehension and abilities. This was apparent in a number of cases in the assessment packages where learners were allowed to choose between questions in some cases. However, the opportunity to choose between questions did not occur often, in most assessment activities learners were not given a choice of questions.

#### **6.5.1.2 The way in which the implementation of a standards-based curriculum influence assessment practices in EFL**

Most of the teachers in the focus group interviews seemed unaware of or unsure of the connection between standards and assessment. They appeared puzzled by the question. This was apparent by their lack of response and apparent confusion. They looked at their colleagues and seemed to search for a response. They appeared ill at ease and their body language (raised eyebrows and puzzled expressions) indicated that they were surprised or confused by the question. One teacher confessed to being confused by the relationship between standards and assessment stating: "It's still very vague". He firmly insisted that the implementation of standards hadn't changed teaching and assessment at all. He maintained that there had not been any real changes in the way that English was taught and assessed. "The whole idea is just semantics", he repeated.

The researcher explained that standards affect assessment and asked the participants how well they felt the Ministry of Education had conveyed the message that assessment changed as standards change. Six teachers shook their heads indicating that they did not think the message was made clear and one of them repeated his earlier comment: "It [assessment] is still vague, it is still very vague. The relationship between standards and assessment is not clear at all."

Only two teachers seemed to be aware that standards affect assessment. They indicated that since students had different abilities they needed to be assessed in different ways. It was pointed out that pupils worked harder {and reached higher standards} when they knew that they were being assessed. Two other teachers indicated that standards did influence writing in particular. They said that as a result of standards there were better and clearer criteria for writing assignments.

It is important to note that although the focus groups indicate that the teachers do not think that standards have influenced assessment, the opposite is true in the assessment packages. The assessment packages show the influence of standards on assessment. Alternative assessment, such as projects, book tasks and literature assignments are apparent in the assessment packages albeit to a limited number.

## **Conclusion**

*Poor understanding of the connection between standard and assessment:* It is clear that most teachers fail to grasp the close connection between standards and assessment. As a result, valuable opportunities to enrich learning are lost. It is clear from section 3.8, *Influence of standards on assessment* that standards encourage, guide, and intensify learning. Since teachers are unaware of the relationship between standards and assessment they are not able to teach or assess effectively. Those who did indicate that they are aware of the influence of standards could offer only limited information on exactly what this influence entails.

The assessment packages confirmed a limited understanding of the connections between standards and learning. Although alternative assessment was present in the

form of projects, performances, book tasks and literature assignments, most of the activities in the assessment packages were comprised of traditional assessment type test items, such as multiple-choice and short fill-in questions. Projects could be found in only six of the assessment packages, literature assignments and/or book tasks and/or oral presentations were in only half of the packages. In four of the packages at least one half or more of the semester grade was based on traditional assessment. In four assessment packages the semester grade was based on over sixty percent of traditional assessment.

*Influence of standards on writing:* A few teachers in the focus group interviews admitted that they saw the influence of standards on writing. They explained that there were better and more precise criterion for writing. This was confirmed by the assessment packages which showed a few examples of rubrics for writing.

### **6.5.1.3 The extent to which EFL teachers still utilise traditional assessment methods and test items?**

One hundred percent of the teachers admitted to still utilising traditional assessment methods and test items. One of them even admitted that traditional assessment is the main assessment method in her classes. The majority of the teachers stated that they often use traditional assessment (multiple-choice; matching; true/false and short answer questions) for checking grammar, vocabulary and reading comprehension. They explained that they used traditional assessment because those types of questions were easy to write and quick to administer and check. One teacher admitted to using traditional assessment as well as other assessment methods, while two others insisted that traditional assessment was not the main assessment that they used.

Reasons given for their reliance on traditional assessment were that it shows what pupils know and reveals their academic standing, and because the structured format of traditional assessment is best for weaker students.

## **Conclusion**

*Reliance on traditional assessment:* The results of the focus group interviews clearly

indicate that teachers rely predominantly on traditional assessment. This is unfortunate since this type of assessment has many disadvantages for both the teacher and the student. It narrows instruction and limits teachers' freedom. It doesn't reflect the principles of cognitive thinking nor does it respect the student or his interests. In traditional assessment the connection between learning and assessment is misunderstood. These disadvantages have been discussed in section 2.5.2, *Criticism of traditional assessment*.

Reliance on traditional assessment was confirmed in the assessment packages. This was apparent in the tests, classwork and homework assignments that were constructed mainly of multiple-choice, fill-ins and short answer questions. Although there was evidence of project work and performance-based tasks, each of these alternative assessment activities were present only once in those packages. The evidence of traditional assessment, on the other hand, was apparent throughout the assessment package.

*Reasons for relying on traditional assessment:* It seems that the time factor is one of the main reasons for relying on traditional assessment. The fact that teachers are of the opinion that traditional assessment benefits weaker students, is once again an indication that learners' abilities determine assessment decisions. The role of time and learners' abilities was also confirmed in 6.5.1.1, *How decisions on the specific type of assessment activities to be used are made*.

#### **6.5.1.4. The extent to which assessment practices make provision for the assessment of all the language skills**

Many teachers mentioned that their assessment practices did not make any provision for the assessment of all the language skills.

Less than half of the teachers admitted that they only assessed the skills that would be tested on the module that their students were taking. They said that consequently, not all the language skills were regularly assessed. Speaking skills, in particular, were often neglected. One of the teachers said: "If the particular skill is not going to be tested on the Bagrut module, we don't teach it or assess it". Another teacher added:

"We simply don't have time to strengthen skills that will not be tested." One teacher complained that the modular system of the Bagrut examinations had resulted in a "modularisation" of skills. He said: "Assessment has become artificial since we are pushed to assess certain skills and ignore others." He explained that he "juggles" his assessment to mirror the grade that he expected students to get on their Bagrut rather than giving them a grade on overall assessment.

One teacher pointed out that she felt that the modular system limited students' opportunities to develop all their language skills and abilities. Many teachers admitted that they didn't often assess speaking skills because it was difficult to assess speaking skills in large classes. Therefore, "speaking skills miss out". Another teacher agreed: "We seem to leave the assessment of speaking skills behind".

One teacher viewed project work as an opportunity to integrate language skills. She gave a lot of assessment in writing and used formative assessment to encourage students to revise and rewrite their work. She involved her students in assessment by providing them with rubrics that guide their writing.

## **Conclusion**

*Failure to assess all language skills:* The focus group interviews revealed that assessment practices failed to make provisions for the assessment of all the language skills. Although class size was mentioned, the main reason for neglecting to assess certain skills was related to the modular Bagrut system. Teachers admitted that since the modular examination was based primarily on traditional assessment, they felt restricted in their teaching and choices of assessment. In addition, to the disadvantages of "teaching to the test" (section 3.6. 3, *Teaching to the test*), the modular Bagrut has a detrimental effect on learning and assessment. The Bagrut interferes with learning and assessment by failing to integrate all the language skills in each module. The modular approach to learning is not in harmony with the principles of cognitivism which promotes the integration of learning skills (section 2.2.2, *Cognitivism*).

The assessment packages confirmed that teachers failed to assess all the language skills. There was only limited evidence of assessment of speaking listening skills and spoken language in the packages. The assessment of writing skills was limited to a small number of projects, book tasks and literature assignments. The assessment packages showed that reading comprehension was often assessed, but mainly by traditional assessment.

*Difficulties assessing speaking skills:* The focus group interviews revealed that many teachers did not assess spoken language because that was not assessed on most modules. In addition, they pointed out the difficulties of assessing students' spoken language, especially in large classes. This attitude was confirmed in the assessment packages where the assessment of spoken language was limited to five examples of oral presentations.

#### **6.5.1.5 The classroom opportunities students have to be assessed by engaging in social interactions in authentic settings and situations**

Very few teachers said that they provided opportunities to assess their students in social interactions in authentic situations. Many teachers considered project work to be the only opportunity for their students to use authentic language. Two teachers said that students engage in social interaction during group work. These activities occurred over a limited period of time.

#### **Conclusion**

*Limited opportunities for social interaction:* The focus group interviews made it clear that the opportunity to engage in social situations in authentic settings was virtually absent from their assessment. They said that the only time when their students were assessed on social interaction was during project work or group work. However, the teachers did not use project work or group work sufficiently. They rarely used group assessment and project work was infrequent. It was clear from section 4.14, *Alternative assessment in EFL in Israel* that language should be assessed in authentic situations with tasks that challenge and encourage learning.

The assessment packages confirmed that there was limited opportunity for social interaction. Only three assessment packages indicated that project work could be done in pairs.

*Problems with assessment:* It is clear from the focus group interviews that teachers have problems understanding assessment. A major element of assessment in EFL is encouraging students to engage in social interaction in the target language for communicative objectives. It is apparent from the focus group interviews that students have been given very little opportunity to do this. Therefore, teachers are experiencing a problem understanding the intention of assessment.

The assessment packages confirm that teachers have problems understanding the objectives of assessment. This was obvious in the absence of assessment activities for social interaction in the assessment packages.

#### **6.5.1.6 The assessment methods and activities that are used on a regular basis**

Very few teachers confirmed that they used alternative assessment methods and activities most of the time. Most teachers agreed that they generally utilised traditional assessment because it most closely mirrored the type of assessment that their students would face on Bagrut. One teacher stated that he used traditional assessment very often because it was the easiest and fastest method of assessment. One teacher said that she used traditional assessment regularly although she said that she used other methods of assessment as well.

One teacher provided additional reasons to explain why she used traditional assessment most of the time. She said that traditional assessment was less demanding than alternative assessment which required greater teacher involvement. In addition, she was concerned that alternative assessment was too difficult for many students since it required them to be adept at negotiating the Internet.

Two teachers admitted to using traditional assessment as well as other assessment which was more suitable for projects and essay writing. A few teachers stated that they used traditional assessment very often because it was the easiest and fastest

method of assessment. One teacher maintained that traditional assessment was the most efficient means of measuring student achievement since "there's a lot of overhead time with alternative assessment". She insists: "The amount of investment in alternative assessment is much more than the amount of pupil gain".

A few teachers said that they believed that traditional assessment showed what pupils knew and revealed their academic standing. One teacher said that he relied on traditional assessment because it assessed Bagrut potential which was what teachers needed to assess. He said that: "Traditional assessment shows the nitty gritty".

Two teachers explained that they used traditional assessment because the structured format of traditional assessment was best for weaker students. A few teachers said they used alternative assessment (for writing and project assignment) as well as traditional assessment.

## **Conclusion**

*Dominance of traditional assessment:* It is apparent that traditional assessment is the assessment that is dominant in most classrooms. The focus group interviews strongly demonstrate that teachers rely on traditional assessment most of the time. The disadvantages of traditional assessment are thoroughly discussed in section 2.5.2, *Criticism of traditional assessment.*

The assessment packages provide irrefutable agreement with the findings of the focus group interviews. Traditional assessment activities are present in every assessment package and dominant most of the assessment in all the packages. Although alternative assessment is visible in the assessment packages, it is an infrequent instrument of assessment.

### **6.5.1.7 The extent to which students are involved in classroom assessment**

The discussions revealed that students are not involved in real classroom assessment. One hundred percent of the teachers admitted that their students were never actively

involved in assessment. One teacher confessed that she didn't think that it was important to involve students in classroom assessment.

Many teachers said that they allowed peer assessment for basic vocabulary work and dictation primarily because it freed them to do other things. One teacher said that she included students in assessment by having them evaluate each other's work. She maintained that students were more attentive and aware when they were involved in assessment. Moreover, they felt more a part of learning when they were actively involved in assessment. Another teacher said that she allowed self-assessment for dictation because it was quick and transmitted the message to her students that she trusted them. A few teachers admitted that they did not consider peer assessment in any final grade

One teacher said that he didn't involve his students in assessment and seemed surprised to hear that some teachers allowed their students to assess each other. He seemed genuinely surprised that students would take part in assessment and asked: "Why do you involve them?"

One teacher was concerned about peer assessment and said that peer assessment led at times to arguments among the students.

## **Conclusion**

*Students not involved in assessment:* It was clear from the discussion that teachers did not see the value of involving their students in assessment. Even in circumstances where peer assessment was permitted, the assessment was not seriously considered by the teacher or the assessment was not on important assessment tasks. The research conducted in this thesis has shown that students who are involved in assessment are more motivated and involved in their learning (section 4.12.4.4, *Self-assessment*; 4.12.4.5, *Peer assessment*). Therefore, teachers fail to maximise their students' learning potential by not including them in assessment.

The assessment packages corroborate the absence of student involvement in assessment. With the exception of a few circumstances in which students completed a

checklist prior to commencing a task or write reflections at the conclusion of a task, they were never involved in assessment. Infrequently students were presented with rubrics beforehand, but they were never involved with creating assessment criterion.

*Problems with assessment:* A problem with assessment is that teachers are unaware of the importance of involving students in assessment. It is clear from section 4.5.6, *Feelings of autonomy* that assessment is most effective when students are involved directly in assessment. The evidence from the focus group interviews and the assessment packages confirms that teachers miss an important opportunity to enhance learning when they fail to include students in assessment.

#### **6.5.1.8 The extent to which assessment activities encourage the improvement of work in progress**

The discussion revealed that formative assessment does not take place regularly. Project work and composition work, however, do encourage the improvement of work in progress. This kind of work, however, is limited.

There was only one teacher who seemed to really understand the significance of formative assessment on learning. She maintained that continuous assessment, rather than a final examination, was essential in order to encourage and nurture learning. She blamed the Bagrut system for restricting learning and assessment.

Another teacher agreed that formative assessment was important, but admitted that it was difficult in large classes. She confessed that she felt guilty for not being able to relate to the continuous assessment needs of individual students and admitted: "I feel like a criminal half of the time".

Most teachers said that they encouraged formative assessment, yet, only in the area of composition and writing where there was the opportunity to correct and revise work

#### **Conclusion**

*Infrequency of formative assessment:* Although assessment encouraged the improvement of work in progress in regard to project work and compositions, formative assessment was non-existent in other areas of learning. It is clear from

section 2.3.3.3, *Formative assessment* and 2.6.1.2, *Continuous assessment and opportunities for improvement* that assessment for learning is achieved through continuous assessment.

The assessment packages confirmed the limited use of formative assessment. The only evidence of formative assessment was in project work and book tasks where students were encouraged to submit a first and then final draft of their work.

*Problems with assessment:* It was clear from the focus group interviews and the assessment packages that formative assessment does not play a significant role in assessment. This is unfortunate since formative assessment, section 2.3.3.3, monitors work in progress and enriches learning. Teachers who do not use formative assessment miss an essential opportunity for effective assessment.

#### **6.5.1.9 Your opinion on the use of criterion-referenced versus norm-referenced assessment**

The discussion revealed that teachers assessed project work and writing assignment with rubrics that were based on criterion. One teacher said that she used criterion-based assessment in order "to measure the steps they {the students} make towards their progress". A few teachers said that they used rubrics for writing. One teacher stated that she provided clear assessment models so pupils knew what weight was given to each area of their activity. The model depended on the level of the class and the level of the assignment.

There was only one teacher who seemed to know a lot about assessment. She pointed out that there were varied levels of assessment and that rubrics were an essential element of assessment. She stated that criterion-referenced assessment was necessary for certain students because they required individualised assessment that needed to be clearly outlined. However, she added that criterion-reference assessment was good for all students.

One teacher said that norm-referenced assessment was necessary since teachers needed to compare students in order to know the level of the class. A few teachers

disagreed. One insisted that norm-referenced assessment wasn't valid since students needed to be assessed according to their own accomplishments. Another said that norm-referenced assessment wasn't relevant since she rarely succeeded in bringing all her students to the same level. An additional teacher agreed that norm-referenced assessment wasn't valid since students needed to be assessed according to their own accomplishments.

## **Conclusion**

*Criterion-referenced assessment:* Criterion-referenced assessment is apparent in project work and writing where students are guided by rubrics. It is obvious from section 2.3.3.7, *Criterion-referenced assessment* that criterion-based assessment contributes to learner productivity.

The assessment packages confirmed that criterion-referenced assessment is very limited in EFL classes. Although a few teachers said that it was important to assess each student according to criterion, others said that it was important to measure a student's achievement in relation to that of others in the class. Evidence of criterion-based assessment was limited in the assessment packages to projects and book tasks where the students were given a set of clear and specific criterion.

*Problems with assessment:* It was obvious from the focus group interviews and the assessment packages that teachers have a limited understanding of the value of criterion-referenced assessment. In many occasions they did assess their students according to criterion-referred assessment. However, criterion-referenced assessment was limited. Teachers apparently fail to understand that criterion-referenced assessment should be used often to help direct students towards successful learning.

### **6.5.1.10 The main reasons why you might be reluctant to use alternative assessment activities**

The focus group interviews very clearly show that EFL teachers are reluctant to use alternative assessment activities.

Many teachers said that the number of students and time considerations discouraged the use of alternative assessment. One teacher admitted that alternative assessment took a lot of time and energy which was in short supply, whereas traditional assessment produced results quickly and easily. Another agreed and said: "Teachers waste valuable time with alternative assessment. We have too much to access and we can do it much quicker and more thoroughly with traditional assessment". One teacher mentioned class size and time restraints as the reasons for her reluctance to use alternative assessment. She complained: "We never have enough time". Another teacher said that she was reluctant to use traditional assessment because it was demanding for the teacher since it involved more teacher involvement.

Fifty percent of the teachers said that they didn't use alternative assessment since that was not the assessment that was on Bagrut. One said: "You need to assess Bagrut knowledge and skills." Another added that since teachers were assessed on Bagrut results they needed to invest their efforts in Bagrut type assessment. The Bagrut also impacts the way that students relate to assessment. One teacher said that students only wanted to engage in activities that were assessed on Bagrut. This, she felt, limited assessment possibilities. She faulted the English Inspectorate for transmitting the message to students that only what is tested on the Bagrut is important. Another agreed and insisted that students are only interested in grades ("that's the bottom line") and that they were not interested in improving their language abilities.

Two teachers admitted that they were hesitant to use alternative assessment because they believed that it was not suitable for all students. They said that they thought that it was good for native speakers because they were best equipped with the skills that were necessary for projects and other performance-based activities. They said that foreign language learners have difficulty with alternative assessment tasks.

## **Conclusion**

*Resistance to alternative assessment:* The focus group interviews clearly indicated that teachers are reluctant to use alternative assessment. Time restrictions and the influence of the Bagrut were cited as the major reason. The challenges of alternative assessment in EFL in Israel are discussed in section 4.16, *The challenges of*

*alternative assessment in EFL in Israel.*

The assessment packages supported the results of the focus group interviews. The assessment in the packages was mostly traditional assessment type activities. This confirmed that teachers are reluctant to use alternative assessment.

*Problems with assessment:* EFL teachers have problems understanding the importance of alternative assessment. As a result, they are reluctant to implement alternative assessment in their classes. The focus group interviews indicated that many teachers didn't feel that it was worth the added time and effort to assess their students in alternative ways. As a result, they missed the opportunity to maximise assessment.

### **6.5.2 Summary of the results of the focus group interviews and assessment packages**

The data that was collected from the focus group interviews shows that alternative assessment has not been fully integrated into the EFL high school classroom. The results of the focus group interviews reveal that most teachers essentially assess their students in the same way they will be assessed on the Bagrut examination. The teachers largely admit that they closely follow the question format of the Bagrut modules in order to prepare their students for the test. These questions are generally multiple-choice and short fill-in questions. Teachers admit that the additional advantage of these types of questions is that they are easy to compose and quick to grade.

Teachers are not aware of the impact that a standards-based curriculum has on assessment. However, they recognise the implementation of clear and descriptive rubrics for assessing writing skills as a result of standards. They fail to understand the important connection between standards, learning and assessment. Only one of the teachers insisted that assessment should not be used to see what a student knew, rather assessment should be used when students were doing well in order to build their self-confidence, to offer encouragement and "to boost their egos". This teacher seems to

be the only one who is aware of the importance of assessment for learning alongside assessment of learning.

The teachers admitted that they do not actively include their students in assessment. Although peer assessment may occur infrequently, teachers said that they did not use the results of peer assessment in their final assessment. Teachers, however, do feel that they involve their students in assessment by giving them rubrics and checklists to follow during assignments. Unfortunately, by not involving students actively in assessment, teachers are failing to acknowledge the role that learners' should assume in assessment (Ministry of Education, Culture and Sport, 2002:11).

The majority of teachers said that they used formative assessment primarily in composition work where they stressed process writing. They maintained that on-going assessment encouraged students to improve their work and helped them learn. One teacher said that he didn't believe that all areas of learning could be improved on an on-going basis. Whereas he said that compositions could be assessed as work in progress, he did not see unseen reading passages in the same light.

Teachers do encourage students to hand in drafts of their written work and to correct, revise and rewrite their compositions. Project work was mentioned as the main area where formative assessment was used. Formative assessment is a principle of assessment that is stressed in *The Assessment Guidelines* which encourage revision and attention to the learning process as well as the final product (Ministry of Education, Culture and Sport, 2002:3, 11, 15). Teachers acknowledged the importance of process writing, however, three teachers mentioned the difficulties they had trying to mark and remark a large number of students' written work.

The teachers in the focus group interviews admitted that they did not assess all language skills. They complained that they felt pressured to assess only the skills that were tested on the Bagrut module that their students were taking. As a result, they didn't spend time teaching or assessing additional skills. Listening skills and oral skills seemed to be ignored by most of the teachers. This was confirmed by the assessment packages. There was limited evidence of assessment activities of listening and oral skills.

Finally, teachers quickly disclosed the main reasons why they were reluctant to use alternative assessment methods and activities. Most of the teachers cited the technical challenges of alternative assessment, such as it being too time consuming and more difficult to use in large classes. Yet, their major reason was that traditional assessment mirrored the assessment their students had on their Bagrut examination. Therefore, they argued that classroom assessment should follow a similar format.

The textual analysis contributed a significant amount of information about the assessment that EFL teachers use in their classrooms. The assessment packages clearly indicated the different assessment methods and activities that teachers regularly used. Traditional assessment is heavily apparent throughout the assessment packages. The tests in all the ten assessment packages were based on traditional assessment items, such as multiple-choice, fill-in and short answer questions. Classwork and homework exercises were similar in most cases. Three teachers gave weekly spelling, vocabulary and grammar quizzes which were formatted on traditional assessment and stressed memorisation and recall.

Alternative assessment activities, such as projects, book tasks and literature assignments were present in the assessment packages of many teachers, but only one example of each of these activities was obvious. It appears that many teachers include alternate assessment, but the occurrence of these activities is very infrequent. Even when there were assessment activities all the tests that were in the packages were largely traditional assessment tests. Projects, book tasks, oral presentations and literature assignments were the only activities that were assessed according to the main criteria of alternative assessment. Most teachers used criterion-referenced assessment to evaluate their students' performance in these areas. Yet, whereas rubrics were present in projects and book tasks, they were rarely present in other types of assessment activities.

## **6.6 SUMMARY**

The empirical study provided a comprehensive insight into the assessment of English foreign language teachers in Israel. Both the focus group interviews and the textual analysis of the assessment packages revealed that EFL high school teachers in Israel

do not sufficiently implement alternative assessment into their teaching and assessment.

The focus group interviews gave evidence that the majority of the teachers admitted that traditional assessment was the main assessment that they use. A few teachers mentioned that they used alternative assessment in project work where they assessed students according to clearly defined criteria. These teachers recognised the new and clearer criteria in writing tasks as a direct result of the implementation of standards. However, teachers insisted that they didn't see the influence of standards on other areas of English foreign language teaching.

The assessment packages showed that teachers did sometimes utilise alternative assessment and that standards did have some influence on assessment. However, evidence from the assessment packages confirmed that alternative assessment assumed a minor role in overall assessment. Traditional assessment is clearly the prevalent assessment method of assessment for EFL teachers in Israel. This was apparent in the assessment packages that teachers submitted and in the focus group interviews.

The assessment packages showed that EFL teachers used alternative assessment methods and activities to a limited degree. This was apparent in many of the assessment packages that contained examples of project work, book tasks and literature assignments. However, there were a limited number of examples of these activities. In all of the assessment packages where these activities were found, they occurred only once, whereas traditional assessment activities took place frequently, usually daily. This is despite a new curriculum and assessment guidelines that have been issued by the Ministry of Education as well as in-service mandatory training to facilitate their implementation.

The focus group interviews indicate that there are several factors which impede the use of alternative assessment. The main factor, according to the teachers, is the disconnection they see between the Bagrut examination and classroom assessment. They are reluctant to use alternative assessment which is disconnected from the traditional type testing items that characterise Bagrut.

In addition, they admit that time constraints and class size make them reluctant to implement alternative assessment. As a result, teachers admit that traditional assessment is their main assessment. A study of the assessment packages confirms the dominance of traditional assessment. Traditional assessment type items comprise most of the assessment in the assessment packages.

The final chapter of this thesis will present the conclusions, implications and recommendations resulting from the empirical study.

## CHAPTER 7

### SUMMARY OF THE STUDY, IMPLICATIONS AND RECOMMENDATIONS OF THE EMPIRICAL STUDY

#### 7.1 INTRODUCTION

In this concluding chapter a general overview of the investigation is given in order to demonstrate that the aims originally expressed have been addressed and achieved. The researcher summarises the study and outlines the limitations of the study, highlights the major conclusions from the study and suggests recommendations including areas for further research. Under limitations, those factors limiting the generalisability of the findings are pointed out and factors that the researcher could not control during the study are discussed.

As stated in section 5.2, *Objectives of the empirical study*, the aim of the empirical study was to identify the activities and methods that English teachers in Israel employ to assess their students. The study also investigated the influence that standards had on assessment, perceptions of assessment and alternative assessment in teaching EFL in Israel.

The findings of the study provide insights into the views, beliefs and practices of the teachers who participated in the empirical study.

#### 7.2 SUMMARY OF THE STUDY

The researcher sets to investigate how EFL teachers in Israel understand the principles underlying assessment and alternative assessment. The main aim of the study was to identify the activities and methods that EFL teachers utilise to assess their students' language ability.

In **chapter 1** the background of the study was discussed. Discussion entailed examining the background of the problem and description of the research methodology. The field of foreign language instruction and assessment has

experienced major developments over the last several decades. As a result the focus of language instruction changed from a concentration on syntax to the teaching of language for communicative purposes in authentic situations.

The standards movement in education was responsible for promoting the creation of standards that clearly define what students should know and be able to do with their knowledge and skills in order to reach academic achievement. In addition, a better understanding of how people obtain and process knowledge also affected the approach to learning. In light of these factors it became obvious that new kinds of assessment were needed. Alternative assessment offered opportunities for students to demonstrate what they were able to do. In 1998 the Ministry of Education in Israel approved a curriculum which included principles and standards for teaching English in all grades. One of the main principles supported by the new curriculum is alternative assessment. Despite measures to improve English foreign language study, many EFL teachers in Israel are experiencing difficulty in implementing standards and alternative assessment in their classes. The intention of this study was to investigate this phenomenon.

**Chapter 1** stated the key research problem: How do EFL teachers incorporate standards and alternative assessment into their assessment practices? In addition, a number of questions were provided to delineate the problem, such as:

- What are the theories underlying assessment?
- How do standards influence assessment?
- What does alternative assessment entail?
- What factors impact the teaching and assessment of English as a foreign language?
- Which assessment activities and methods do English foreign language teachers employ?

The remaining chapters comprise the bulk of the thesis.

**Chapters 2, 3 and 4** constituted the literature study. In **chapter 2** the researcher investigated the role of assessment in learning. A comprehensive study of the theories such as: behaviorism, cognitivism and constructivism that underlie assessment shed

light on how the mind accepts, and processes knowledge. This led to a deeper perception of what constitutes real learning and resulted in the understanding that real learning is born out of learning experiences that reflect real-life situations. Cognitive and constructivist theories placed the learner in the center of the learning process. Student-centered learning places responsibility on the learner and motivates him or her to assume responsibility for academic achievements. Consequently, traditional assessment which had been a popular means of measuring rote learning and recollection of isolated bits of information needed to be replaced with alternative assessment which allows students to use their knowledge and skills in real-life settings and situations. Alternative assessment challenges students to use higher order thinking in order to construct meaning and solve problems.

In addition, alternative assessment has brought the student to the center of the learning process, which has a monumental effect on how students react and relate to education. Student-centered learning has been proved to be an important element of achievement. Students who are directly involved in learning and assessment show greater interest, motivation and ultimate success.

Students need the opportunity to direct and manage their own learning. Assessment activities, such as portfolio reflection, can help them develop an awareness of what they have learned. Self-assessment and peer and group assessment encourage students to take responsibility for their own academic success. The transition from traditional to alternative assessment has paved the way for greater academic achievement.

**Chapter 3** discussed the implementation of standards in education and its impact on assessment practices. The standards movement was the result of dissatisfaction with existing school systems and the belief that not all children have the benefit of equitable learning opportunities. Standards intended to clarify what all students needed to accomplish in order to reach the same benchmarks. Clear and reasonable standards were developed that specify what students are able to do with their knowledge and skills to promote learning. Standards are the best chance of assuring that there are equal learning opportunities for all students.

Standards, themselves, do not guarantee academic success. Several factors must be considered in order for standards to be effective. Standards have to be specific and very clearly defined. The number of standards has to be limited so that they can be taught and the standards should be reachable at some level for all students despite differences in their backgrounds. Lastly, teachers must be trained with the knowledge and skills they need in order to teach according to the standards.

Standards have a direct influence on assessment since assessment should demonstrate the level to which students meet the standards. Therefore, assessment methods and activities have to be altered to reflect the principles underlying the standards and they need to be modified so that they are suitable for different students and their diverse needs. This means that assessment should consider that not all students work at the same pace nor are all students able to excel in all tasks. Allowing students a choice of activities and a flexible timetable will maximise the benefits of standards and assessment.

Although there have been considerable objections to adopting standards-based curricula, educators in countries around the world have embraced standards as a promising solution to the problems that plague their educational systems. Israel has followed the examples of other foreign language countries and adopted a curriculum based on standards. In Israel, the movement to standards has implemented several changes in the approach to assessment. *The New Curriculum* has transformed English teaching in Israel by focusing attention on assessment and placing the student at the center of the learning process. The implementation of a standards-based curriculum is apparent in the attention to criterion-referenced assessment and formative assessment which encourages motivation and supports assessment for learning.

**Chapter 4** discussed the factors that influence the teaching and assessment of English as a foreign language in Israel. The dominant role of English in the 21<sup>st</sup> century as well as globalisation has made English even more important than in the past. The emphasis on learning a language for communicative purposes has affected the way that foreign language is taught. Cognitive theories of learning have influenced foreign language teaching, learning and assessment in Israel. This is apparent in alternative

assessment activities, such as projects, presentations and portfolios. These activities encourage higher order thinking and train students to be independent learners.

Alternative assessment shifts the focus of assessment to task-based learning and peer and group assessment. Teachers are encouraged to implement alternative assessment such as projects, portfolios and presentations in their classes. These activities develop critical thinking and train students to use their knowledge and skills to construct new meaning. Modifications have recently been made in the matriculation examinations to incorporate higher level thinking and improved instruction. The Bagrut has been changed from one three hour examination to a series of three test modules. In addition, there have been changes in the type of questions that comprise the tests.

**Chapter 5** focused on the research design and methodology of the empirical study. The researcher gave the rationale for choosing a qualitative research plan and explained the two research procedures, textual analysis and focus group interviews, which would comprise the bulk of the study. The criteria used to examine and judge the influence of standards on assessment and the questions for the focus group interviews were introduced. The chapter explained the procedures for collecting, sorting, analysing and presenting the research data.

**Chapter 6** presented the findings of the empirical study. The analysis of the assessment packages and the focus group interviews revealed essential information about EFL teachers' assessment beliefs and practices. This information has important implications regarding assessment in the study of English as a foreign language. The results have particular significance in Israel, but the findings shed light on similar situations in other countries.

The data that was collected from the twelve participants of the focus group interviews and the textual analysis of ten assessment packages reveals that alternative assessment has not been fully integrated by EFL high school teachers in Israel. The empirical investigation disclosed that, despite the implementation of a new curriculum that is based on standards and that promotes alternative assessment, most teachers rely primarily on assessment that is similar to Bagrut assessment.

### 7.3 SYNOPSIS OF FINDINGS

The findings of this qualitative investigation are presented and integrated with prior research on alternative assessment in foreign language teaching as reviewed in chapters two to four.

- Despite alternative assessment guidelines that were distributed in 2002 (section 4.14.1, *Assessment Guidelines for the English Curriculum*), much of the assessment activities in the assessment packages that were submitted for analysis, still relies heavily on traditional assessment activities (section 1.7.2, *Traditional assessment*), such as written tests that primarily consist of multiple-choice, fill-in and other types of closed questions. Although project work, literature assignments and book tasks do offer students opportunities to engage in authentic activities (section 4.12, *Authentic assessment in EFL*), and encourage autonomous learning and self-assessment (section 4.12.4.4, *Self-assessment*), these types of activities are not used as often as traditional assessment activities.
- The majority of the teachers are reluctant to use alternative assessment. This can be linked to several factors. They regard alternative assessment as too time consuming and difficult to use in large classes. However, the main reason discussed was the Bagrut (section 4.15, *Matriculation examination - Bagrut*). More than half of the teachers believe that since the Bagrut is comprised primarily of traditional assessment type items, they should assess in a similar manner. This was confirmed by the textual analysis of the assessment packages which showed that the tests in all ten packages were modeled on the Bagrut modular that students would be taking that semester. In addition, quizzes, classwork and homework included similar types of questions, such as, multiple-choice, fill-in and short answer questions that mirror the Bagrut modules. The detrimental effect of the new modular Bagrut on classroom assessment could clearly be seen and teachers are extremely dissatisfied with this system.

- The modularisation of the Bagrut has caused most of the teachers to choose assessment that prepares their students for Bagrut matriculation (section 4.15.1, *Background about Bagrut*). Consequently, they "teach to the test" (section 3.6.3, *Teaching to the test*) which has resulted in assessment that is, as has already been mentioned, mostly restricted to traditional assessment type test items such as multiple-choice, fill-in and short answer questions. Teachers stated that the additional advantage of these types of questions is that they are easy to write and quick to grade.
- Even though all teachers in Israel have been required to take at least one in-service course to familiarise themselves with assessment theory, many teachers seem unaware of the principles underlying assessment (section 2.2, *Theories underlying assessment practices*) and do not incorporate them into their assessments. This could explain why alternative assessment is not an essential part of their teaching. This might also explain why self and group assessment (section 4.12.4.4, *Self-assessment*; section 4.12.4.6, *Group assessment*) is virtually non-existent in their classrooms. Their assessment packages clearly indicate that classroom assessment is seldom created with these principles in mind.
- Although the Israel Ministry of Education issued a standards-based curriculum in 2001 (section 3.10.1, *The New Curriculum*), the majority of teachers are unfamiliar with the standards that are outlined in the new curriculum and there is little attention to standards in their assessment.
- Most teachers don't believe that standards have had a major impact on assessment (section 3.8, *Influence of standards on assessment*). The study did, however, show that teachers recognise the influence of standards on writing. A few teachers in the focus group interviews pointed out the valuable use of rubrics in writing (section 2.3.4.2, *Rubrics*), yet they don't seem to see the impact of standards on other areas. Examples of rubrics for projects were evident in many of the assessment packages.

- The majority of teachers have a misconception regarding assessment. Teachers associate assessment primarily with testing (section 2.3.2, *Testing and assessment*) and fail to understand the impact that assessment can have on learning (section 2.6.1.7, *Improvement of learning*). More specifically, they do not view alternative assessment as an important means of promoting and encouraging language learning.
- More than half of the teachers have many problems with assessment and feel frustrated and confused. They fault the modularised system of the Bagrut, in which the four language skills appear in separate modules, for preventing them from integrating all of the language skills (section 4.12.4.1, *Projects* and narrowing the scope of teaching and assessment. In addition, they experience difficulties dealing with their students' unwillingness to be involved in alternative assessment activities that they don't perceive to be directly related to the Bagrut. Some teachers also feel that it is unfair to assess students in a way that is opposed to the formal assessment that they will have for matriculation. These views give evidence to the powerful influence of teaching to the test (section 3.6.3, *Teaching to the test*).
- Most teachers do not actively include their students in assessment (section 4.5.6, *Feelings of autonomy*), but a few teachers indicated that they do present their students with rubrics for project presentation prior to writing so that they know the criterion that will be used to assess their work (section 4.15.4, *Projects in EFL in Israel*). Most teachers stated that they use formative assessment (section 2.3.3.3, *Formative assessment*) in projects and in composition. Students are encouraged to submit a first draft of their writing which they can correct, revise and rewrite. Most teachers agreed that project work was the area in which students were assessed while their work was in progress as well as at completion.
- Although alternative assessment is clearly absent from most assessment in the classroom, more than half of the teachers do occasionally involve their students in projects and creative literature assignments and book tasks. These

activities are the areas in which teachers do use criterion-referenced assessment (section 2.3.3.7, *Criterion-referenced assessment*), checklists and rubrics, and formative as well as summative assessment (section 2.3.4.1, *Checklists*; section 2.3.4.2, *Rubrics*; section 2.3.3.3, *Formative assessment*; section 2.3.3.4, *Summative assessment*).

The hypotheses that were presented in section 5.2, *Objectives of the empirical study*, have all been confirmed in the empirical study.

## 7.4 RECOMMENDATIONS

This section presents the recommendations based on research questions that formed the study as well as conclusions arrived at from the literature review and the empirical investigation.

- Teachers should gain a better understanding of the implications of alternative assessment on teaching prior to becoming teachers i.e. during collegiate studies as preparation for the BEd in English.

Although each of the teachers who participated in the empirical study had taken at least one in-service course on the new curriculum and alternative assessment, it is clear from the evidence that one course is insufficient to prepare them sufficiently for the challenges of incorporating alternative assessment. It is of the utmost importance that they receive comprehensive, on-going training in the principles underlying standards and assessment as well as serious training in creating and implementing alternative assessment methods and activities. The Ministry of Education should invest in more and better training for teachers. Just as assessment must be on-going, teacher training in this area must also be continuous.

- Teachers' opinions, thoughts and concerns about assessment should be heard and considered by the English Inspectorate of the Ministry of Education.

It is apparent that teachers are not only confused by assessment, but are also frustrated. Efforts must be made to listen to these teachers and address their fears and frustrations. Teachers are the guardians of effective assessment. Assessment for

learning cannot take place if teachers fail to understand the principles of assessment or if they are uncertain about assessment. The English Inspectorate of the Ministry of Education must dispatch representatives to all schools and encourage teachers to address their reservations and uncertainties about assessment. This can be done through school meetings with the English teaching supervisory board or with visits by teacher trainers during school hours. Study days with workshops that focus essentially on assessment should also be organised.

- A review of the Bagrut modular system.

The intense dissatisfaction with the new Bagrut modular system makes it clear that the modularised Bagrut warrants to be reevaluated. A thorough investigation by the English Inspectorate of the Ministry of Education is needed which may ultimately call for changes to be implemented. In light of the strong criticism it seems that a return to the three hour Bagrut examination that integrated assessment of all the language skills is a more efficient measurement of EFL achievement. Clearly, the current attention to the modularised examinations sacrifices classroom assessment and discourages teachers.

Teachers need support. They are unfamiliar with the principles of assessment and have difficulty implementing alternative assessment. They regard alternative assessment as time-consuming. Therefore, they should be provided with help marking assignments or be allocated fewer classes so that they can face the challenges presented by alternative assessment.

- Implementation of quality supervision.

The quality of teacher supervision is seriously lacking. The teachers expressed their confusion regarding assessment. As a result, they continue to rely on traditional assessment as the primary assessment in their classes although the *NBA Handbook*, the *Principles and Standards for Learning English as a Foreign Language in all Grades* and *Assessment Guidelines for the New Curriculum* clearly require that teachers implement alternate assessment. The fact that traditional assessment remains the main assessment would make it seem that teachers are not being supervised nor monitored. Supervision should go hand in hand with support.

- A change in teachers' attitude.

The teachers' attitude regarding assessment appears to be a significant obstacle to the implementation of alternative assessment. Most teachers admit that they use traditional assessment because it is easier and less time consuming. They admit that they don't think that alternative assessment warrants the added time, effort and involvement that it requires. A major obstacle is their belief that they must "teach to the test" so that their students will succeed on the Bagrut. In addition, the modularisation of the Bagrut interferes with the assessment of all the language skills.

If Bagrut is changed, and if teachers get the necessary support so that they understand the importance of assessment for learning, it will hopefully contribute to a more positive attitude.

The continuous use of traditional assessment practices raises issues that are a cause of great concern to anyone interested in quality education. The results should be of particular interest to the English Inspectorate of the Ministry of Education for the following reasons:

### **Recommendations for further research**

The researcher recommends an investigation that will encompass the assessment practices of a greater sampling of teachers in all grades.

- A comparative study of assessment in classes where the students do not take a Bagrut module at the end of the semester (9<sup>th</sup> and 10<sup>th</sup> grade) with classes that do (11<sup>th</sup> and 12<sup>th</sup> grade).

This investigation will show whether assessment is more integrated when there is no Bagrut module that semester.

- A replication of this study in other EFL contexts and other settings in order to add more validity to the current conclusions.

Study should be repeated under a larger population and an analysis of the findings in such a study to see which teachers (e.g. more experienced teachers of smaller classes, of homogeneous classes, etc.) implement or do not implement alternative assessment.

### **7.5 CONTRIBUTIONS OF THE RESEARCH**

The objective of this study was to investigate teachers' understanding of alternative assessment and find out how teachers incorporate alternative assessment in their EFL classrooms. The study, conducted through two focus group interviews and the textual analysis of ten assessment packages provides a genuine look into the actual teaching practices of a sampling of EFL high schools teachers. Since the researcher is unaware of any similar studies that have been conducted in Israel, she is confident that this research has illuminated the subject and sowed the seeds for additional investigation.

The researcher does not presume to suggest that this current study is in any way complete. Despite efforts to provide in-depth and comprehensive research on the topic, it is apparent that there is a necessity for further investigation. The convenience sampling of 11<sup>th</sup> grade high school teachers limits the results of the study. However, the use of a qualitative research design yielded important valid and important information that gives a reliable presentation of the subject. The empirical study reveals several conclusions that might inform teacher education in Israel especially in the area of alternative assessment in the teaching of English as a foreign language in Israel. The current findings can be helpful in the planning and design of more comprehensive studies that may have future ramifications.

### **7.6 LIMITATIONS OF THE STUDY**

While the contribution of this study to the field of assessment in English as a foreign language cannot be underestimated, the universal applicability and generalisability of the findings is limited by the following two factors:

- The scale of the study
- Learners' opinions were not investigated

The study was limited to a convenience sampling of 11<sup>th</sup> grade EFL high school teachers. Twelve teachers participated in the focus group interviews and ten assessment packages comprised the textual analysis. The researcher acknowledges that this is a relatively small sampling of the population.

The study did not investigate the learners' opinions. Clearly, an understanding of students' attitude towards assessment would have contributed to the scope of the investigation. However, this was not the focus of the study. The study focused on how EFL teachers understand alternative assessment and how they incorporate alternative assessment in their EFL classrooms.

Notwithstanding the above, findings from this study provide pertinent pointers to a number of crucial educational issues warranting careful consideration by and attention from individuals who are involved in determining educational policies, in teacher training or in teaching. It should be taken into consideration that the teachers who participated in the study were mostly trained by the same trainers from the English Inspectorate of the Ministry of Education. This gives credibility to the findings and conclusions arrived at. Chances are high that any observed limitations in the handling of the assessment practices based on standards and any views held by teachers who formed the two focus groups and whose assessment packages were analysed, are likely to be a general reflection of the situation prevailing throughout the country.

## **7.7 CONCLUDING REMARKS**

The results of this study suggests that the English Inspectorate of the Ministry of Education in Israel has not been successful in implementing alternative assessment in the teaching of English as a foreign language in Israel. Despite the creation of a new curriculum based on standards, the state of assessment in the country has remained relatively static. Assuming the 'population' studied is representative of the majority of the teachers in the country it appears that traditional assessment is the predominant assessment method albeit limited examples of alternative assessment do occur. The evidence of this is inconclusive from the empirical study that was undertaken.

Considering the impact of these findings, it is very important for those responsible for

the success of English foreign language education in Israel, mainly the English Inspectorate of the Ministry of Education, to examine the existing situation.

While the scope of the findings of this study might seem small, the magnitude should be interpreted in terms of the overall impact that these findings might have on the study of English as a foreign language in Israel as well as on the educational system in the country.

Therefore, an important message of this investigation is that educational policies must be studied in the field. Although several documents have been published that introduce and advocate revolutionary changes in assessment in the country, it cannot be assumed that these changes have been implemented successfully in the field.

Finally, this study has enlightened the researcher on the subject of this thesis. The results of this investigation have already proved valuable in her teaching and clearly her students are the beneficiaries of her greater understanding of this subject. The study has whetted her appetite for further investigation into the area of assessment in foreign language teaching.

## Bibliography

- Airasian, A. 1997. *Classroom assessment* (3<sup>rd</sup> ed.). New York: McGraw-Hill.
- Airasian, P. & Walsh, M. 1997. Constructivist cautions. *Phi Delta Kappan*, 78(6): 444-449.
- Alderson, J.C. & Banerjee, J. 2001. Language testing and assessment. *Language Teaching*, 34(4): 213-236.
- Alderson, J.C. & Wall, D. 1993. Does washback exist? *Applied Linguistics*, 14(2): 115-129.
- Alexander, F. 1993. National standards: a new conventional wisdom. *Educational Leadership*, 50(5): 9-10.
- Allan, D. 2000. Testing and assessment in ELT. *ETAI Forum* XI(3): 26-28.
- Alvermann, D.E. & Phelps, S.F. 1994. *Content reading and literacy: succeeding in today's diverse classrooms*. Boston: Allyn and Bacon.
- American Council on the Teaching of Foreign Languages (ACTFL). 1996. Standards for Foreign Language Learning: Preparing for the 21<sup>st</sup> Century. <http://www.actfl.org/files/public/execsumm.pdf> (accessed on 2003-04-14).
- American Council on the Teaching of Foreign Languages (ACTFL). 1998. ACTFL Performance Guidelines for K-12 Learners. <http://www.actfl.org/i4a/pages/index/cfm?paged=3327i> (accessed on 2003-04-14).
- Anderson, G. & Arsenault, N. 1998. *Fundamentals of educational research* (2<sup>nd</sup> ed.). London: Falmer.
- Armstrong, T. 1994. *Multiple intelligences in the classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Arter, J. 1995. Portfolios for assessment and instruction. *ERIC Digest*. [http://www.ed.gov/databases/ERIC\\_Digests/ed388890.html](http://www.ed.gov/databases/ERIC_Digests/ed388890.html) (accessed on 2004-06-19).
- Arter, J. 2000. Rubrics, scoring guides and performance criteria: classroom tools for assessing and improving student learning. *ERIC Digest*. <http://www.eric.ed.gov/ed446100.html> (accessed on 2005-03-03).
- Assessment Reform Group. 2002. Assessment for learning. [www.assessment-reform-group.org.uk](http://www.assessment-reform-group.org.uk) (accessed on 2006-12-12).
- Assinder, W. 1991. Peer teaching, peer learning. *ELT Journal*, 45(3): 218-228.
- Atkins, M.J. 1993. Theories of learning and multimedia applications: an overview. *Research Papers in Education*, 8(2): 251-271.

- Auerbach, E. 1993. Reexamining English only in the ESL classroom. *TESOL Quarterly*, 27(1): 9-32.
- Bachman, L.F. 1990. *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Baker, E.L., O'Neil, H.F. & Linn, R.L. 1991. Policy and validity prospects for performance-based assessment. Paper presented at the annual meeting of the American Psychological Association, 24 August 1991.
- Barth, P. & Mitchell, R. 1992. *Smart start: elementary education for the 21<sup>st</sup> century*. Colorado: North American.
- Bell, T.R. 2005. Behaviors and attitudes of effect. *Foreign Language Annals*, 38(2): 259-270.
- Berg, B.L. 1998. *Qualitative methods for the social sciences*. Boston, MA: Allyn and Bacon.
- Biggs, J.B. 1995. Assumptions underlying new approaches to educational assessment. *Curriculum Forum*, 4(2): 1-22.
- Black, P. & Wiliam, D. 1998. Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 80(2): 139-149.
- Boomer, G. 1992. Negotiating the curriculum. In: G. Boomer, N. Lester, C. Onore, & J. Cook (eds). *Negotiating the curriculum*. London: The Falmer Press. pp. 4-14.
- Boud, D. & Brew, A. 1995. *Enhancing learning through self-assessment*. London: Kogan Page.
- Brandt, R. 1993. On teaching and understanding: a conversation with Howard Gardner. *Educational Leadership*, 50(7): 4-7.
- Brandt, R. 1995. Overview: what to do with those new standards. *Educational Leadership*, 52(6): 5.
- Broadfoot, P., Murphy, R. & Torrance, H. (eds) 1990. *Changing educational assessment: international perspectives and trends*. London: Routledge.
- Brown, H.D. 1994. *Teaching by principles: an interactive approach to language pedagogy*. Englewood Cliffs, NJ: Prentice Hall Regents.
- Brown, H.D. 2001. *Teaching by principles*. New York: Longman.
- Brown, H.D. 2004. *Language assessment: principles and classroom practices*. White Plains: Pearson Education.
- Brown, J.D. & Hudson, T. 1999. The alternatives in language assessment. *TESOL Quarterly*, 32(4): 653-673.

- Brown, S. 1990. Assessment: a changing practice. In: T. Horton (ed.). *Assessment debates*. London: Hodder & Stoughton. pp. 5-11.
- Brualdi, AC. 1996. Multiple intelligences: Gardner's theory. *ERIC Digest*. [http://www.ed.gov/databases/ERIC\\_Digests/ed410226.html](http://www.ed.gov/databases/ERIC_Digests/ed410226.html) (accessed on 2002-03-15).
- Brutt-Griffler, J. 1998. Conceptual questions in English as a world language: taking up an issue. *World Englishes*, 17(3): 381-392.
- Buhsmer, K. 2000. Tele-Internet focus groups: an alternative to traditional focus groups. *Quirk's Marketing Review*. [http://www.quirks.com/articles/article\\_print.asp?arg\\_articleid=557](http://www.quirks.com/articles/article_print.asp?arg_articleid=557) (accessed on 2006-03-25).
- Burden, P. 2000. The use of the students' mother tongue in monolingual English "conversation" classes at Japan University. *The English Language Teacher*, 24(6): 5-10.
- Burt, M. & Keenan, F. 1995. Adult ESL learner assessment: purposes and tools. *ERIC Digest*. [http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content\\_storage](http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage) (accessed on 2002-01-21).
- Cates, K. 1990. Teaching for a better world: global issues in language education. *The Language Teacher*, 14(5): 3-5.
- Chamot, A.U. 2004. Issues in language learning strategy research and teaching. *Electronic Journal of Foreign Language Teaching*, 1(1): 14-26. <http://e-flt.nus.edu.sg/vIn2004/chamot.pdf> (accessed on 2005-12-20).
- Cody, W.S. 2000. Designing an effective performance task for the classroom. *English Teachers' Journal*, 53: 15-23.
- Cohen, D. 1990. Reshaping the standards agenda. In: P. Broadfoot, R. Murphy & H. Torrance (eds). *Changing educational assessment: international perspectives and trends*. London: Routledge: 32-51.
- Cohen, D. 1995. What standards for national standards? *Phi Delta Kappan*, 76(10): 751-758.
- Cohen, L. & Morrison, M.L. 2000. *Research methods in education* (5<sup>th</sup> ed.). London: Rutledge Falmer.
- Collins Cobuild English Language Dictionary. 1992. Sinclair, J (ed.). London: HarperCollins.
- Coltrane, B. 2002. New standards for foreign language teaching: a vision for the future. *The Eric Review*, 26(1): 6-8.

Commission on Teaching Standards for School Mathematics. 1991. Professional standards for teaching mathematics. Reston, VA: The National Council of Teachers of Mathematics.

Cotton, J. 1995. *The theory of assessment*. London: Kogan Page.

Covington, M. 1996. The myth of intensification. *Educational Researcher*, 25(8): 24-26.

Crandall, J. & Tucker, G.R. 1990. Content-based instruction in second and foreign languages. In: A. Padilla, H.H. Fairchild, & C. Valadez (eds). *Foreign language education: issues and strategies*. Newbury Park, CA: Sage. pp. 187-200.

Cresswell, J.W. 1994. *Research design: qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Cresswell, J.W. 1998. *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks, CA: Sage.

Cresswell, J.W. 2005. *Educational research: planning, conducting and evaluating quantitative and qualitative research* (2<sup>nd</sup> ed.). Upper Saddle River, NJ: Pearson.

Crookes G. & Schmidt, R.W. 1991. Motivation: reopening the research agenda. *Language Learning*, 41: 469-512.

Crystal, D. 1997. *English as a global language*. Cambridge: Cambridge University Press.

Dakwar, L. 2003. Concerns about projects. *ETAI Forum*, 14(2): 56.

Darling-Hammond, L., Aness, J. & Falk, B. 1995. *Authentic assessment in action - studies in schools and students at work*. New York: Columbia University Teachers College.

Denzin, N.K. & Lincoln, Y.S. 1994. *Handbook of qualitative research*. London: Sage.

Diegmueller, K. 1995. Running out of steam. Struggling for standards: an *Education Week* special report, 14(29): 4-6. Washington DC: Education Week.

Dietal, R.J., Herman, J.L. & Knuth, R.A. 1991. *What does research say about assessment?* NCREL: Oak Brook.

Dornyei, Z. 2001. *Teaching and researching motivation*. England: Pearson Education.

Dyer, B. & Bushell, B. 1996. World issues or a global perspective. [www.jalt-publications.org/tit/files/96/nov/global.htm](http://www.jalt-publications.org/tit/files/96/nov/global.htm) (accessed on 2005-10-19).

Eisner, E.W. 1991a. What really counts in schools. *Educational Leadership*, 48(5): 10-11, 14-17.

Eisner, E.W. 1991b. *The enlightened eye: qualitative inquiry and the enhancement of educational practice*. New York, NY: Macmillan.

Eisner, E.W. 1993. Why standards may not improve schools. *Educational Leadership*, 50(5): 22-23.

Eisner, E.W. 1999. The use and limits of performance assessment. *Kappan Professional Journal*, 80(9): 658-661.

Ellis, R. 2000. Task-based research and language pedagogy. *Language Teaching Research*, 4(3): 193-220.

English Teachers' Network in Israel (ETNI). [www.etni.org](http://www.etni.org) (accessed on 2006-09-20).

Eskey, D.E. 1997. Syllabus design in content-based instruction. In: M.A. Snow & D.A. Brinton (eds). *The content-based classroom: Perspectives on integrating language and content*. White Plains, NY: Longman. 132-141.

Estaire, S. & Zanon, J. 1994. *Planning classwork: a task-based approach*. Oxford: Macmillan Heinemann.

Evers, W.M. 2001. Standards and accountability. In: T. Moe (ed.). *A Primer on American's Schools*. Stanford, CA: Hoover Institutions: 205-247.

Fisher, L., Evans, M. & Esch, E. 2004. Computer-mediated communication: promoting learner autonomy and intellectual understanding at the secondary level. *Language Learning Journal*, 30: 50-58.

Fraser, W.J., Loubser, C.P. & Van Rooy, M.P. 1993. *Didactics for the undergraduate student*. Durban: Butterworth.

Frazier, D.M. & Paulson, F.L. 1992. How portfolios motivate reluctant writers. *Educational Leadership*, 49(8): 62-65.

Frye, S. 1991. Communicating the next message of reform through the professional standards for teaching mathematics. *ERIC/SMEACH Mathematics Education Digest*. <http://www.ericdigests.org/1992-5/reform.htm> (accessed on 2003-03-20).

Gagne, E.D., Yekovich, C.W. & Yekovich, F.R. 1993. *The cognitive psychology of school learning* (2<sup>nd</sup> ed.). New York: HarperCollins College Publishers.

Gandal, M. 1995. Not all standards are created equal. *Educational Leadership*, 52(6): 16-21.

Gandal, M. & Vranek, J. 2001. Standards: here today, here tomorrow. *Educational Leadership*, 59(1): 6-13.

Gardner, H 1987. *The mind's new science: a history of the cognitive revolution*. New York: Basic Books.

- Gardner, H. 1993. *Multiple intelligences: the theory in practice*. New York: Basic Books.
- Gardner, H. 1994. "Reinventing our schools: a conversation with Howard Gardner" – a video conversation.
- Gardner, H. 1995. Reflections on multiple intelligences: myth and message. *Phi Delta Kappan*, 77(3): 200-209.
- Gardner, R.C. & MacIntyre, P.D. 1992. A student's contribution to second language learning. Part I: Cognitive variables. *Language Teacher*, 25(4): 211-220.
- Gardner, R.C. & MacIntyre, P.D. 1993. A student's contribution to second language learning. Part II: Affective variables. *Language Teacher*, 26(1): 1-11.
- Genesee, F. & Upshur, J.A. 1996. *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.
- Ghosn, I.K. 1997. Teaching EFL to multiple intelligences. *ERIC Digest*. [http://www.ed.gov/databases/ERIC\\_Digests/ed438713.html](http://www.ed.gov/databases/ERIC_Digests/ed438713.html) (accessed on 2006-02-02).
- Gibbs, C.V. 1994. *Beyond testing: towards a theory of educational assessment*. London: Farmer Press.
- Glaser, R. & Silver, E. 1994. *Assessment, testing, and instruction: retrospect and prospect*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- Good, T.L. & Brophy, J.E. 1994. *Looking in classrooms* (6<sup>th</sup> ed.). New York: HarperCollins.
- Grady, M.P. 1998. *Qualitative and action research: a practitioner handbook*. Bloomington, Indiana: Phi Delta Kappan Educational Foundation.
- Graham, S. 1997. *Effective language learning*. Clevedon, Avon: Multilingual Matters.
- Gratz, D.B. 2000. High standards for whom? *Phi Delta Kappan*, 81(9): 681-687.
- Gravett, S. & Geysler, H. 2004. *Teaching and learning in higher education*. Pretoria: Van Schaik.
- Gronlund, N.E. & Linn, R.L. 1990. *Measurement and evaluation in teaching* (6<sup>th</sup> ed.). New York: Macmillan.
- Hamayan, E.V. 1995. Approaches to alternative assessment. *Annual Review of Applied Linguistics*, 15: 212-226.
- Hamp-Lyons, L. & Condon, W. 1999. *Assessing the portfolio: principles for practice, theory, and research*. Cresskill, NJ: Hampton.

Hancock, C.R. 1994. Alternate assessment and second language study: what and why? *ERIC Digest*. [http://www.ed.gov/databases/ERIC\\_Digests/ed 376695.html](http://www.ed.gov/databases/ERIC_Digests/ed 376695.html) (accessed on 2002-01-29).

Hanna, G.K. & Dettmer, P.A. 2004. *Assessment for effective teaching - using context-adaptive planning*. Boston: Pearson.

Harbord, J. 1992. The use of the mother tongue in the classroom. *EFL Journal*, 46(4): 30-55.

Harnischfeger, A. 1996. *Fads or Reform? The standards movement in the United States*. Chicago: Northwest University.

Harris, J. 2000. Re-visioning the boundaries of learning theory in the assessment of prior experimental learning (APEL). Paper presented at SCUTREA. 3-5 July 2000. University of Nottingham. <http://www.leeds.ac.uk/educol/documents/00001448.htm> (accessed on 2002-05-05).

Harris, M. 1997. Self-assessment of language learning in formal settings. *ELT Journal*, 51(1): 12-20.

Hasman, A. 2000. The role of English in the 21<sup>st</sup> century. *Forum*, 38(1): 2-5.

Hatch, J.A. 2002. Accountability shovedown: resisting the standards movement in early childhood education. *Phi Delta Kappan*, 83(6): 457-462.

Hatch, T. 1997. Getting specific about multiple intelligences. *Educational Leadership*, 54(6): 26-29.

Hawkins, E.W. 1999. Foreign language study and language awareness. *Language Awareness*, 8 (3&4): 124-142.

Hawkins, M.R. (ed.). 2004. *Language learning and teacher education*. Clevedon, England: Multilingual Matters.

Herbert, E. 2001. *The power of portfolios. What children can teach us about learning and assessment*. San Francisco: Jossey-Bass.

Herman, J. 1992. Developing performance based assessment tasks. <http://www.lacoe.edu/pdc/second/assessment.html> (accessed on 2000-04-17).

Herman, J.L., Aschbacher, P.R., & Winters, L. 1992. *A practical guide to alternate assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Hibbard, K.M. 1996. *A teacher's guide to performance-based learning and assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

- Horsburgh, D. 2003. Evaluation of qualitative research. *Journal of Clinical Nursing*, 12(2): 307-312.
- Huerta-Macias, A. 1995. Alternate assessment: response to commonly asked questions. *TESOL Journal*, 5(1): 8-11.
- Irie, K. 2003. What do we know about the language learning motivation of university students in Japan? *JALT Journal*, 25(6): 86-100.
- Jamentz, K. 2001. *The instructional demands of standards-based reform*. Washington, DC: American Federation of Teachers.
- Johnson, S.D. 1995. Will our research hold up under scrutiny? *Journal of Industrial Teacher Education*, 32(3): 3-6.
- Johnson, R.B. & Christensen, L.B. 2004. *Educational Research*. Boston: Allyn and Bacon.
- Katz, A. 1999. Standards-based assessment for ESL students. *ERIC/CLL News Bulletin*, 22(2): 1-3.
- Kemp, J. & Toperoff, D. 1998. Guidelines for portfolio assessment in teaching. Jerusalem: Ministry of Education, Culture and Sport.
- Kendall, J.S. & Marzano, R.J. 1996. *Content knowledge: a compendium of standards and benchmarks for K12 educators*. Aurora, CO: Mid-continent Research for Education and Learning.
- Kilfoil, W.R. & Van der Walt, C. 1997. *Learn 2 teach – English language teaching in a multilingual context*. Pretoria: J.L.van Schaik.
- Kleiber, P. 2004. Focus groups: more than a method of qualitative inquiry. In: K. deMarrai. & S.D. Lapan (eds). *Foundations for research – Methods of inquiry in education and the social sciences*. Mahwah, NJ: Lawrence Erlbaum. 88-102.
- Kluth, P. & Straut, D. 2001. Standards for diverse learners. *Educational Leadership*, 59(1): 43-46.
- Kohonen, V. 1999. Authentic assessment in affective foreign language education In: J. Arnold (ed.). *Affect in language learning*. Cambridge: Cambridge University Press: 279-294.
- Kramsch, C. 1992. Interactive discourse in small and large groups. In: W. Rivers (ed.). *Interactive language teaching*. New York: Cambridge University Press: 17-30.
- Krieger, D. 2005. Teaching ESL versus EFL – principles and practices. *Forum*, 43(2): 8-16.
- Krippendorff, K. 2004a. Measuring the reliability of qualitative text analysis data. *Quality & Quantity*, 38: 787-800.

Krippendorff, K. 2004b. Reliability in content analysis: some common misconceptions and recommendations. *Human Communication Research*, 30(3): 411-433.

Krueger, R.A. 1994. *Focus groups* (2<sup>nd</sup> ed.). Thousand Oaks: Sage.

Krueger, R.A. 2002. Designing and conducting focus group interviews. [www.shadac.org/collecting/collateral/Handout%20October2002pdf](http://www.shadac.org/collecting/collateral/Handout%20October2002pdf) (accessed on 2005-04-27).

Larsen-Freeman, D. 2000. *Techniques and principles in language teaching* (2<sup>nd</sup> ed.). Oxford: Oxford University Press.

Leloup, J. & Ponterio, R. 1998. Meeting the national standards: now what do I do? *ERIC Digest ED425657*. [www.ericdigests.org/1999-3/standards.htm](http://www.ericdigests.org/1999-3/standards.htm) (accessed on 2005-03-20).

Lessard-Clouston, M. 1997. Language learning strategies: an overview for L2 teachers. *Internet TESL Journal* 3(12). <http://iteslj.org?Articles?Lessard-Clouston-Strategy.html> (accessed on 2003-11-24).

Lewis, A.C. 1995. An overview of the standards movement. *Phi Delta Kappan*, 76(10): 744-750.

Lier, L.V. 1996. *Interaction in the language curriculum*. London: Longman.

Linn, R.L., Baker, E.L. & Dunbar, S.B. 1991. Complex performance-based assessment: expectations and validation criterion. *Educational Researcher*, 20(8): 15-21.

Linn, R.L. & Gronlund, N.E. 2000. *Measurement and assessment in teaching* (8<sup>th</sup> ed.). Upper Saddle River, NJ: Merrill.

Little, D. 1991. *Learner autonomy 1: definition, issues and problems*. London: Longman.

Littlewood, W. 2004. The task-based approach: some questions and suggestions. *ELT Journal*, 58(4): 319-326.

Lofland, J. & Lofland, L. 1995. *Analysing social settings: a guide to qualitative observation and analysis*. New York: Wadsworth.

Looney, J. 2005. England: implementing formative assessment in a high stakes environment. In: *Formative assessment: improving learning in secondary classrooms*. Paris: OECD. pp. 129-147.

Looney J., Laneve, C. & Moscato, M.T. 2005. Italy: a system in transition. In: *Formative assessment: improving learning in secondary classrooms*. Paris: OECD: 163-175.

- Lopez-Reyna, N.A. & Bay, M. 1997. Enriching assessment: using varied assessments for diverse learners. *Teaching Exceptional Children*, 29(4): 33-37.
- Loyd, B.H. & Loyd, D.E. 1995. Kindergarten through grade 12 standards: a philosophy of grading. In: G.D. Phye (ed.). *Handbook of classroom assessment, learning, adjustment and achievement*. San Diego: Academic Press: 481-490.
- Macmillan, J.H. 2001. *Classroom assessment: principles and practice for effective instruction* (2<sup>nd</sup> ed.). Boston: Allyn and Bacon: 481-489.
- Lucas, E. 2005. Conflict resolution and global issues in EFL classes. *ETAI Forum*, XVII (1): 42-44.
- Macaro, E. 1997. *Target language, collaborative learning and autonomy*. Clevedon: Multilingual Matters.
- MacIntyre, P. & Gardner, R. 1991. Methods and results in the study of anxiety and language learning: a review of the literature. *Language Learning*, 41(1): 85-117.
- MacIntyre, P.D., Noels, K.A., & Clement, R. 1997. Biases in self-ratings of second language proficiency: the role of language anxiety. *Language Learning*, 47(2): 265-287.
- Maehr, M.L. & Maehr, J.M. 1996. Schools aren't as good as they used to be: they never were. *Educational Researcher*, 25(8): 21-29.
- Maeroff, G.I. 1991. Assessing alternative assessment. *Phi Delta Kappan*, 73(4): 273-281.
- Maley, A. 1992. Global issues in ELT. *Practical English Teaching*, 13(2): 73.
- Mann, G., Shemesh, R. & Shlayer, J. 2001. Projects at work. *English Teachers' Journal*, 54: 18-39.
- Mantle-Bromley, C. 1995. Positive attitudes and realistic beliefs: links to proficiency. *Modern Language Journal*, 79(2): 372-386.
- Maoz, S. 2000. Alternative assessment. *English Teachers' Journal*, 53:91-94.
- Marshall, C. & Rosman, G.B. 1994. *Designing qualitative research*. Thousand Oaks: Sage.
- Marzano, R.J. & Kendal, J.S. 1996. Designing standards- based districts, schools, and classrooms. Alexandria, VA: Association for Supervision & Curriculum Development.
- Marzano, R.J. & Kendall, J.S. 1997. The fall and rise of standards-based education. A National Association of School Boards of Education (NASBE) Issues In Brief. Alexandria, VA. [http://www.mcrel.org/PDF/Standards/5962IR\\_FallAndRise.pdf](http://www.mcrel.org/PDF/Standards/5962IR_FallAndRise.pdf) (accessed on 2002-03-22).

- Marzano, R.J. & Kendall, J.S. 1998. *Awash in a sea of standards*.  
[www.mcreo.org/PDF/Standards/59821R\\_AwashInASea.pdf](http://www.mcreo.org/PDF/Standards/59821R_AwashInASea.pdf) (accessed on 2002-03-22).
- Mayer, R.E. 2005. The failure of educational research to impact educational practices: six obstacles to educational reform. In: G.D. Phye, D.H. Robinson & J. Levine (eds). *Empirical methods for evaluating educational interventions*. Burlington, MA: Elsevier Academic Press: 67-81.
- Mayring, P. 2000. Qualitative content analysis. *Forum Qualitative Research*, 1(2).  
<http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm> (accessed on 2006-03-06).
- McDonough, J.J. & McDonough, S. 1997. *Research methods for English language teachers*. London: Arnold.
- McKay, S.L. 2002. *Teaching English as an international language: rethinking goals and approaches*. Oxford: Oxford University Press.
- McNamara, M. & Deane, D. 1995. Self-assessment activities: toward autonomy in language learning. *TESOL Journal*, 5(1): 17-21.
- McTighe, J. 1997. What happens between assessments? *Education Leadership*, 54(4): 6-12.
- McTighe, J & O'Connor, K. 2005. Seven practices for effective learning. *Educational Leadership*, 63(3): 10-17.
- McTighe, J. & Thomas, R.S. 2003. Backward design for forward action. *Educational Leadership*, 60(5): 52-55.
- Merton, R.K., Fiske, M. & Kendall, P.L. 1990. *The focused interview: a manual of problems and procedures* (2<sup>nd</sup> ed.). London: Collier MacMillan.
- Medgyes, P. 1994. *The non-native teacher*. London: Macmillan.
- Messick, S. 1994. Standards-based score interpretation: establishing valid grounds for valid inferences. Princeton, NJ: Educational Testing Service.
- Messick, S. 1996. Validity and washback in language testing. *Language Testing*, 13(3): 241-256.
- Met, M. 1991. Language learning through content: learning content through language. *Foreign Language Annals*, 24(4): 281-295.
- Mitchell, R. & Myles, F. 1998. *Second language learning theories*. London: Arnold.
- Morgan, D.L. 1997. *Focus groups as qualitative research* (2<sup>nd</sup> ed.). London: Sage.

- Moya, S. & O'Malley, M. 1994. A portfolio assessment model for ESL. *The Journal of Educational Issues of Language minority students*, 13: 15-36.  
<http://www.ncela.gwu.edu/pubs/jeilms/vol13/portfo13.htm> (accessed on 2005-10-10).
- Murphy, T. 1995. Tests: learning through negotiated interaction. *TESOL*, 4(2): 12-16.
- Ministry of Education, Culture and Sport. 1995. *Policy for language education in Israel (in Hebrew)*. Jerusalem: Office of the Director General.
- Ministry of Education, Culture and Sport. 1995, 1996. *Policy on language education in Israeli schools*. Jerusalem: Ministry of Education, Culture and Sport.
- Ministry of Education, Culture and Sport, 1998. *The new curriculum for the teaching of English in Israel – principles and standards for learning English as a foreign language in Israeli schools*. Jerusalem: Ministry of Education, Culture and Sport.
- Ministry of Education, Culture and Sport. 2001. *English curriculum: principles and standards for learning English as a foreign language for all grades*. Jerusalem: Ministry of Education, Culture and Sport.
- Ministry of Education, Culture and Sport. 2002. *Assessment guidelines for the English Curriculum (2001)*. Jerusalem: Ministry of Education Pedagogical Secretariat, English Inspectorate.
- Ministry of Education, Culture and Sport. 2003. *New Bagrut assessment handbook (NBA)*. Jerusalem: Ministry of Education, Culture and Sport.
- Ministry of Education, Culture and Sport. 2004. *Professional standards for English teachers: knowledge and performance*. Jerusalem: Ministry of Education, Culture and Sport.
- Mitchell, R. & Myles, F. 1998. *Second language learning theories*. London: Arnold.
- Murphy, R. & Torrance, H. 1990. The need for change. In: T. Horton (ed.). *Assessment debates*. London: Hodder & Stoughton:12-18.
- Neill, M. 2003. The danger of testing. *Educational Leadership*, 60(5): 43-46.
- Niyogi, N. 1995. *The intersection of instruction and assessment: the classroom*. Princeton, NJ: Educational Testing Service.
- Nunan, D. 1996. Learners strategy training in the classroom: an action research study. *TESOL Journal*, 6(1): 35-41.
- Nunan, D. 1999. *Second language teaching and learning*. Boston: Heinle & Heinle.
- Nunes, A. 2004. Portfolios in the EFL classroom: disclosing an informed practice. *ELT Journal*, 58(4): 327-335.

Olsen, D. 1997. Symposium on the implications of new perspectives in cognition and constructivism for social studies, social science, and environmental education. Paper presented at the Fourth International Conference in Studies of Society and the Environment. The University of Sydney, Australia. 6-9 July 1997.

Olson, L. 2000. Worries of a standards "backlash" grow. *Eduweek*, 19(30).  
www.eduweek.org/ew/ewsrtory.cfm?slug=30stand.h19 (accessed on 2005-03-20).

Oosterhof, A. 1999. *Developing and using classroom assessments* (2<sup>nd</sup> ed.). Upper Saddle River: Merrill Prentice Hall.

O'Malley, J.M. & Chamot, A.U. 1990. *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

O'Malley, M. & Valdez Pierce, L. 1996. *Authentic assessment for English language learners*. New York: Addison-Wesley.

O'Neil, J. 1993. Can national standards make a difference? *Educational Leadership*, 50(5): 4-8.

Orkin, J. 2004. Projects and pupil power. *ETAI Forum* XV(3): 36-38.

Oxford, R.L. 1990a. *Language learning strategies: what every teacher should know*. New York: Newbury House.

Oxford, R.L. 1990b. Styles, strategies, and aptitude: connections for language learning. In: T.S. Parry & C.W. Stansfield (eds). *Language aptitude reconsidered*. Englewood Cliffs, NJ: Prentice Hall: 67-125.

Oxford, R.L. 1992. Language learning strategies in a nutshell: update and ESL suggestions. *TESOL Journal*, 2(2): 18-22.

Oxford, R. & McGroarty, M. 1990. Language learning strategies; An introduction and two related studies. In: A.M. Padilla, H.H. Fairchild & C.M. Valadez (ed.). *Foreign language education - issues and strategies*. Newbury Park, CA: Sage. pp. 56-74.

Padilla, A.M. & Sung, H. 1990. Information processing and fl learning. In: A.M. Padilla, H.H. Fairchild & C.M. Valadez (ed.). *Foreign language education - issues and strategies*. Newbury Park, CA: Sage: 45-55.

Patton, M.Q. 1990. *Qualitative evaluation and research methods* (2<sup>nd</sup> ed.). Newbury Park, CA: Sage.

Perkins, D. 1992. *Smart schools: from training memories to using minds*. New York: The Free Press.

Philips, D. 1995. The good, the bad and the ugly: the many faces of constructivism. *Education Researcher*, 24(7): 5-12.

- Pierce, L.V & O'Malley, J.M. 1992. *Performance and portfolio assessment for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Piller, I. 2002. Passing for a native speaker: identity and success in second language learning. *Journal of Sociolinguistics*, 4(2): 179-206.
- Pintrich, P.R. & Schunk, D.H. 1996. *Motivation in education: theory, research and applications*. Englewood Cliffs, NJ: Prentice-Hall.
- Plavin, S. 1998. Questions you always wanted to ask about portfolio but were afraid to ask. *ETAI Forum*, IX(3): 44-46.
- Popay, J., Rogers, A. & Williams, G. 1998. Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research*, 8(3): 341-351.
- Popham, J.W. 2001. Standardised achievement tests: misnamed and misleading. *Education Week*, 21(3): 46.
- Popham, W.J. 2003. *Test better, teach better: the instructional role of assessment*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).
- Popham, W.J. 2004a. A game without winners. *Educational Leadership*, 62(3): 46-50.
- Popham, W.J. 2004b. Teaching to the test – an expression to eliminate. *Educational Leadership*, 62(3): 82-83.
- Powell, R.A. & Single, H.M. 1996. Focus groups. *International Journal of Quality in Health Care*, 8(5): 499-504.
- Prodromou, L. 1999. The great globe itself: Shakespeare, EFL and global issues. *IATEFL Issues*, October-November: 2-5.
- Prodromou, L. 2000. From mother tongue to other tongue. <http://www.thracenet.gr/bridges/bridges5> (accessed on 2003-02-12).
- Puchler, N. & Fried, K. 2001. *Learning to teach modern foreign language in the secondary school*. London: Routledge.
- Punch, K.F. 1998. *Introduction to social research*. London: Sage.
- Ratnesar, R. 1999. The homework ate my family. *Time*, 25 January 2007: 55-63.
- Ravitch, D. 1995. *National standards in American education: a citizen's guide*. Washington: Brookings Institute.

- Ravitch, D. 1996. The state of standards.  
<http://www.edexcellence.net/institute/publication/publication.cfm?id=223> (accessed on 2003-01-05).
- Ravitch, D. 1999. Higher, but hollow, academic standards.  
<http://www.edexcellence.net/institute/publication/publication.cfm?id=219>
- Rea-Dickins, P. & Rixon, S. 1997. The assessment of young learners of English as a foreign language. In: C. Clapman & D. Corson (eds). *Encyclopedia of language and education, language testing and assessment* (vol. 7). Dordrecht: Kluwer: 151-161.
- Rechtman, A.S. 1999. New curriculum – a change or a holiday. *Trends*, 8: 23-25.
- Reigeluth, C. 1997. Educational standards: To standardise or to customise learning? *Phi Delta Kappan*, 79(3): 202-206.
- REL Network. 1999. So many standards, so little time.  
[www.relnetwork.org/news/news26.html](http://www.relnetwork.org/news/news26.html) (accessed on 2002-03-04).
- Resnick, L. & Nolan, K. 1995. Where in the world are world-class standards? *Educational Leadership*, 52(6): 6-11.
- Resnick, L. & Resnick, D.P. 1992. Assessing the thinking curriculum: new tools for educational reform. In: B.R. Gifford & M.C. O'Connor (eds). *Changing assessments: alternative views of aptitude, achievement and instruction*. Boston: Kluwer: 37-75.
- Rigg, P. 1991. Whole language in TESOL. *TESOL Quarterly*, 25(3): 521-542.
- Rogers, T.B. 1995. *The psychological testing enterprise*. Pacific Grove, CA: Brooks/Cole.
- Ruddock, G. 1995. England and Wales. In: *Performance standards in education – in search of quality*. Paris: Organization for Economic Cooperation and Development (OECD): 79-102.
- Rust, C. 2002. Purposes and principles of assessment. In: Learning and Teaching Briefing Papers Series. Oxford Centre for Staff and Learning development OCSLD. [http://www.brookes.ac.uk/services/ocsd/2\\_learnth/briefing\\_papers/p\\_p\\_assessment.pdf](http://www.brookes.ac.uk/services/ocsd/2_learnth/briefing_papers/p_p_assessment.pdf) (accessed on 2004-01-21).
- Sabers, D. & Sabers, D. 1996. Conceptualizing, measuring, and implementing higher (high or hire) standards. *Educational Researcher*, 25(8): 9-21.
- Sachs, J. 2003. Teacher professional standards: controlling or developing teaching? *Teachers and Teaching*, 9(2): 175-186.
- Sadler, D.R. 1998. Formative assessment: revisiting the territory. *Assessment in Education: Principles, Policy and Practice*, 5:77-84.

- Sadler, P. & Good, E. 2006. The impact of self- and peer-grading on student learning. *Educational Assessment*, 11(1): 1-31.
- Sandrock, P. 1997. State standards: connecting a national vision to local implementation. *An ACTFL Newsletter*: 7-13.
- Scherer, M. 2001. How and why standards can improve student achievement - A conversation with Robert J. Marzano. *Educational Leadership*, 59(1): 14-18.
- Schmoker, M. & Marzano, R.J. 1999. Realising the promise of standards-based education. *Educational Leadership*, 56(6): 17-21.
- Schweers, W, 1999. Using L1 in the L2 classroom. *Forum*, 37(2): 6-9.
- Sebba, J. & Maxwell, G. 2005. Queensland, Australia: an outcomes-based curriculum In: *Formative assessment: improving learning in secondary classrooms*. Paris: OECD: 191-203.
- Shaaban, K. 2001. Assessment of young learners. *English Teaching Forum*, 39(4): 16-22.
- Sharkey, J. 1994-5. Helping students become better learners. *TESOL Journal*, 4(2): 18-23.
- Shavelson, R.J. & Baxter, G.P. 1992. Measuring the quality of mathematics and science education. In: P. Vedder (ed.). *Measuring the quality of education*. Lisse: Swets and Zeitlinger: 35-54.
- Shepard, L.A. 2000. The role of assessment in a learning culture. *Educational Review*, 29(7): 4-14.
- Sheppard, K. & Stoller, F.L. 1995. Guidelines for the integration of student projects into ESP classrooms. *Forum*, 33(2).  
<http://exchanges.stste.gov/forum/vols/vol33/no2/p10.htm> (accessed on 2006-04-17).
- Shlayer, J. 2000. Using student portfolios as an assessment tool in second language learning. *English Teachers' Journal*, 53: 34-38.
- Shohamy, E. 1992. Beyond proficiency testing: a diagnostic feedback testing model for assessing foreign language learning. *Modern Language Journal*, 76(4): 513-521.
- Shohamy, E. 2000. Inside the "black box" of classroom language tests. *English Teachers' Journal*, 53: 54-60.
- Shohamy, E., Donitsa-Schmidts, & Forman, I. 1996. Test impact revisited: washback effect over time. *Language Testing*, 13(3): 298-317.
- Sieborger, R. & Macintosh, H. 2004. *Transforming assessment: a guide for South African teachers* (2<sup>nd</sup> ed.). Cape Town: Juta.

- Sifakas, N.C. & Sougari, A. 2003. Facing the globalisation challenge in the realm of English language teaching. *Language and Education*, 17(1): 59-71.
- Sirotnik, K.A. & Kimball, K. 1999. Standards for standards-based accountability systems. *Phi Delta Kappan International*.  
<http://www.pdkintl.org/kappan/ksir9911.htm> (accessed on 2003-01-22).
- Skehan, P. 2002. A non-marginal role for tasks. *ELT Journal* 56(3): 289-295.
- Sliwika, A., Fushell, M. & Gauthier, M. 2005. Canada: encouraging the use of summative data for formative purposes. In: *Formative assessment: improving learning in secondary classrooms*. Paris: OECD: 97-115.
- Smelser, N.J. & Baltes, P.B. (eds). 2001. *International encyclopedia of the social and behavioral sciences* (vol. 4). Amsterdam Elsevier Science Ltd.
- Smith, K. 1997. Action research on the use of portfolio for assessing foreign learners. *English Teachers' Journal*, 50: 31-37.
- Smith, K. 2002. Children's rights, assessment, and the digital portfolio: is there a common denominator? *ETAI Forum*, XIII(2): 18-21.
- Smith, K. 2004. Standards - accept or reject? *ETAI Forum*, XV(1): 83-85.
- Smith, M.L. & Fey, P. 2000. Validity and accountability in high stakes testing. *Journal of Teacher Education*, 51(5): 334-344.
- Solomon, J. 1997. Language teachers align curricula with standards: preliminary results of a national survey. *CAL*, 21(1).  
<http://www.cal.org/resources/News/199709/9709Main.html> (accessed on 2006-02-20).
- Solomon, P.G. 2002. *The assessment bridge - positive ways to link tests to learning, standards and curriculum improvement*. Thousand Oaks, CA: Corwin Press.
- Spolsky, B. 1992. Diagnostic testing revisited. In: E. Shohamy & R.A. Walton (eds). *Language assessment and feedback: testing and other strategies*. Dubuque: Kendall/Hunt: 29-39.
- Spolsky, B. 2002. Teaching globalised English in Israel in the 21<sup>st</sup> century: rights and responsibilities. A talk given at the ETAI Winter Seminar at the Arab College, Haifa.
- Spolsky, B. & Shohamy, E. 1996. National profiles of languages in education. Israel: Language Policy Research Center Publications.  
<http://www.biu.ac.il/hu/lprc/lprcprof.htm> (accessed on 2003-08-10).

Stahl, R.J. 1996. Cognitive psychology and constructivist psychology: implications for and applications to thinking and learning in social studies. Paper presented at the annual meeting of the College of the University faculty assembly of the National Council for the Social Studies, Washington, D.C.

Steiner, J. 2000. Implications of a standards-based curriculum for the teaching-learning-assessment process. *English Teachers' Journal*, 53: 9-14.

Stern, H.H. 1992. *Issues and options in language teaching*. Oxford: Oxford University Press.

Stewart, D.W. & Shamdasani, P.N. 1990. *Focus groups: theory and practice*. London: Sage.

Stiggins, R.J. 2001. The unfilled promise of classroom assessment. *Educational Measurement: Issues and Practices*, 20(3): 5-15.

Stiggins, R.J. 2002. Assessment crisis: the absence of assessment for learning. *Phi Delta Kappan*, 83(10): 758-765.

Stiggins, R.J. 2005. *Student-involved assessment for learners* (4<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson Education.

Stoller, F.L. 2002. Project work: A means to promote language content. *English Teachers' Journal*, 55: 9-17.

Strauss, A. & Corbin, J. 1990. *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Swain, M. 2000. The output hypothesis and beyond: mediating acquisition through collaborative dialogue. In: J.P. Lantolf (ed.). *Sociocultural theory and second language learning*. Oxford: Oxford University Press: 97-114.

Sword, W. 1999. Accounting for presence of self: reflections on doing qualitative research. *Qualitative Health Research*, 9(2): 270-278.

Tang, J. 2002. Using L1 in the English classroom. *Forum*, 40(1): 36-43.

Tannenbaum, J. 1996. Practical ideas on alternative assessment for ESL students. <http://www.cal.org/resources/Digest/tannen01.html> (accessed on 2003-04-17).

TESOL. 1997. *ESL standards for pre-K-12 students*. Alexandria, VA: TESOL.

Thompson, L. 2000. Foreign language assessment: 30 years of evolution and changes. *ERIC/CLL News Bulletin*, 23(2): 3-5.

Thompson, S. 2001. The authentic standards movement and its evil twin. *Phi Delta Kappan*, 82(5): 358-362.

- Todd, R.W. 2002. Using self-assessment for evaluation. *English Teaching Forum*, 49(1): 16-19.
- Torrance, H. (ed.). 1995. *Evaluating authentic assessment*. Buckingham: Open University Press.
- Tremblay, P.F. & Gardner, R.C. 1995. Expanding the motivation construct in language learning. *Modern Language Journal*, 79(4): 508-518.
- Tucker, M. & Coddling, J. 1998. Raising our standards for the standards movement. *Education Week on the Web*. [www.edweek.com/ew/vol-17/23tucker.h17](http://www.edweek.com/ew/vol-17/23tucker.h17) (accessed on 2004-01-21).
- Tudor, I. 1996. *Learner-centredness as language education*. Cambridge: Cambridge University Press.
- Ur, P. 1996. *A course in language teaching: practice and theory*. Cambridge: Cambridge University Press.
- Ur, P. 2003. The very last word: shattering the myths II. *ETAI Forum*, XIV(2): 63-66.
- Van Rooyen, M. & Prinsloo, F. 2002. *Outcomes-based assessment*. Cambridge: Cambridge University Press.
- Vann, R. & Abraham, R. 1990. Strategies of unsuccessful language learners. *TESOL Quarterly*, 24(2): 177-198.
- Vaughn, S., Shay, J. & Sinagub, J. 1996. *Group interviews in education and psychology*. Newbury Park, CA: Sage.
- Vavrus, L. 1990. Put portfolios to the test. *Instructor*, 100(1): 48-53.
- Verma, G.K. & Mallick, K. 1999. *Researching education: perspectives and techniques*. London: Falmer Press.
- Voogt, J. & Kasurinen, H. 2005. Finland: emphasising development instead of competition and comparison. In: *Formative assessment: Improving learning in secondary classrooms*. Paris: OECD: 149-162.
- Walker, J.C. & Evers, C.W. 1999. Research in education: epistemological issues. In: J.P. Keeves & G. Lakomski (eds). *Issues in educational research*. Amsterdam: Pergamon. pp 40-56.
- Weber, R. 1990. *Basic content analysis* (2<sup>nd</sup> ed.). Newbury Park, CA: Sage.
- Wenden, A. 1998. *Learner strategies for learner autonomy*. Great Britain: Prentice Hall.
- Widdowson, H.G. 1994. The ownership of English. *TESOL Quarterly*, 28(2): 377-389.

- Wiersma, W. 1995. *Research methods in education*. Boston: Allyn and Bacon.
- Wiggins, G. 1991. Standards, not standardization: evoking quality student work. *Educational Leadership*, 48(5): 18-25.
- Wiggins, G. 1992. Creating tests worth taking. *Educational Leadership*, 49(8): 26-33.
- Wiggins, G.P. 1993. *Assessing student performance: exploring the purposes and limits of testing*. San Francisco: Jossey-Bass.
- Wiggins, G. & McTighe, J. 2000. *Understanding by design*. Upper Saddle River: Merrill/Prentice Hall.
- Wiggins, G. & McTighe, J. 2006. Examining the teaching life. *Educational Leadership*, 63(6): 26-29.
- Williams, M. & Burden, R. 1997. *Psychology for language teachers: a social constructivist approach*. Cambridge: Cambridge University Press.
- Willis, J. 1996. *A framework for task-based learning*. London: Longman.
- Worthen, B.R. 1993. Critical issues that will determine the future of alternative assessment. *Phi Delta Kappan*, 74(60): 444-454.
- Young, D. 1991. Creating a low-anxiety classroom environment: what does language anxiety research suggest. *The Modern Language Journal*, 75(4): 426-439.

## Appendix 1: Teachers' Assessment Packages

### Textual Analysis: Criteria used to determine whether alternative measures were incorporated

Criterion	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6	Teacher 7	Teacher 8	Teacher 9	Teacher 10
1 Traditional questions such as multiple choice, matching, true and false, fill in, essays and paragraphs	*	**	*	***	*	*	*	***	*	**
2 Memorisation or rote learning of isolated bits of information are required and encouraged	—	**	—	***	—	—	—	***	—	***
3 Students are assessed on oral activity	*	—	—	—	—	—	*	—	—	—
4 Students are actively engaged in assessment	*	*	—	—	—	—	*	—	—	*
5 Opportunities for self-assessment and peer assessment are provided	*	*	*	—	—	—	*	—	—	*
6 Provision is made to assess all the language skills	*	*	—	—	*	—	*	—	—	*
7 Assessment encourages varied mediums for assessment, such as journals, projects, diaries, student logs, etc.	*	—	—	—	*	—	*	—	—	*
8 Formative assessment encourages the improvement of work in progress	**	*	*	—	*	—	*	*	—	*
9 Opportunities for criterion-based assessment are provided	**	*	—	—	*	—	*	—	*	**
10 Assessment exercises are based on the goals of instruction and class material	***	**	**	**	**	**	**	**	**	***

**Key: Evidence of the criterion in one semester (4 lessons weekly over 5 months = approx. 80 lessons**

- Minimal or no evidence in assessment package

\* Evident in 30% of the lessons

\*\* Evident in 50% of the lessons

\*\*\* Evident in 75% or more of the lessons

\*\*\*\* Evident in every lesson

## Appendix 2: Teachers' Assessment Packages

### Textual Analysis: Criteria used to judge the influence of standards

Criterion	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6	Teacher 7	Teacher 8	Teacher 9	Teacher 10
1 Tasks demonstrate what students are required to know and able to do	**	—	**	—	*	*	*	—	*	—
2 Task is an exemplary demonstration worthy of emulation	**	—	*	—	*	*	*	—	*	—
3 Tasks evidence vital basic knowledge in particular subject area rather than rote learning	**	—	*	—	*	*	*	—	*	—
4 Task encourages cognitive growth	*	—	*	—	*	*	*	—	*	—
5 Tasks maximises students' learning potential	*	—	—	—	—	—	—	—	—	—
6 Tasks are specific and clear and can be easily assessed	**	*	*	**	**	*	*	*	*	*
7 Tasks require students to use knowledge and skills in authentic situations	*	—	*	—	*	*	*	—	*	*
8 Task requirements are flexible allowing for student diversity	—	—	—	—	—	—	—	—	—	—
9 Tasks allow for students to use their multiple intelligences	—	—	—	—	—	—	—	—	—	—

**Key: Evidence of the criterion in one semester (4 lessons weekly over 5 months = approx. 80 lessons)**

- Minimal or no evidence in assessment package

\* Evident in 30% of the lessons

\*\* Evident in 50% of the lessons

\*\*\* Evident in 75% or more of the lessons

\*\*\*\* Evident in every lesson

**Appendix 3: The percentage distribution of the semester grade**

**Assessment Packages**

<b>Semester Grade (100%)</b>	<b>Teacher 1</b>	<b>Teacher 2</b>	<b>Teacher 3</b>	<b>Teacher 4</b>	<b>Teacher 5</b>	<b>Teacher 6</b>	<b>Teacher 7</b>	<b>Teacher 8</b>	<b>Teacher 9</b>	<b>Teacher 10</b>
<b>Tests</b>	30%	50%	40%	60%	30%	40%	30%	70%	30%	60%
<b>Literature</b>	20%	20%	–	15%	15%	20%	–	–	–	–
<b>Projects</b>	30%	20%	20%	–	20%	20%	–	–	30%	–
<b>Book Tasks</b>	–	–	20%	–	10%	–	20%	30%	–	20%
<b>Oral Presentation</b>	20%	–	–	–	10%	10%	–	–	20%	10%
<b>Listening</b>	–	–	10%	10%	–	10%	–	–	–	–
<b>Classwork &amp; Homework</b>	–	10%	10%	15%	15%	–	50%	–	20%	10%