

Háskóli Íslands

Hugvísindasvið

Enska

Assessing English Learning in Iceland

*A Study of Assessment Methods Used
in Icelandic Primary and Secondary Schools*

Ritgerð til M.Paed.-prófs

Birna Margrét Arnþórsdóttir

Kt.: 130661-5649

Leiðbeinandi: Birna Arnbjörnsdóttir

Maí 2009

Acknowledgements

I would like to give my sincere thanks to all the people who, in some way or other, have been of assistance to me in writing this thesis. To all the participating English teachers, for their help in making this study a reality. To my supervisor, Professor Birna Arnbjörnsdóttir, for her help and encouragement. To all my former students, for giving me countless reasons to study further. Last, but not least, to my dear family, my husband and three children, for their support and patience through my years of studies, where I, at times, neglected them somewhat.

Table of Contents

Acknowledgements	2
Table of Contents	3
List of Tables	6
Abstract	7
1 Introduction	8
1.1 The Purpose of the Study	9
1.2 Organization of the Thesis	10
2 Literature Review	12
2.1 Definitions of Terms	12
2.1.1 Assessment and Measurement	12
2.1.2 Classification	13
2.1.3 Summative Assessment	14
2.1.4 Formative Assessment	15
2.1.5 Validity and Reliability	16
2.2 Learning Theories and Assessment	17
2.3 Classroom-based Assessment	20
2.4 Purposes and Methods	21
2.4.1 Tests	22
2.4.2 Performance-based Assessment	24
2.4.3 Portfolio Assessment	24
2.4.4 Self-assessment	26
2.4.5 Peer-assessment	27
2.5 Putting Formative Assessment into Practice	28
2.6 Concerns Regarding Language Assessment	30

2.7 Impact of Assessment	33
2.8 Icelandic National Curriculum Guidelines	34
2.8.1 The Guidelines for The Primary Level	35
2.8.2 The Guidelines for the Secondary Level	37
2.9 A Review of Previous Research on Classroom-based Assessment	38
2.9.1 International Research	38
2.9.2 Icelandic Research	42
3 The Study	45
3.1 The Subjects	45
3.2 The Questionnaires	46
4 The Results	48
4.1 Teachers' Education and Experience	48
4.2 Summative Assessment	50
4.2.1 Assessment Methods	50
4.2.2 Types of Tasks Used	57
4.2.3 Language Skills Assessed	61
4.3 Formative Assessment	68
4.4 Presentation of Assessment Results	70
5 Discussion	72
5.1 Differences Regarding Teachers' Education and Experience.....	72
5.2 Written Tests	73
5.3 Alternative Assessment Methods, In-class Effort and Homework	75
5.4 Variety in Assessment Methods	78
5.5 Assessment of All Language Skills	79
5.6 Formative Assessment	80

5.7 Presentation of Assessment Results	82
6 Conclusion	84
References	87
Appendix	93
The Questionnaire for Primary Level Teachers – English Version	93
The Questionnaire for Primary Level Teachers – Icelandic Version	99
The Questionnaire for Secondary Level Teachers – English Version	104
The Questionnaire for Secondary Level Teachers – Icelandic Version	104

List of Tables

Table 1. Secondary level teachers' education and English teaching experience	49
Table 2. Primary level teachers' education and English teaching experience	49
Table 3. Assessment methods used for summative assessment	50
Table 4. Types of written tests used	52
Table 5. Connection between assessment methods and teachers' education and experience	56
Table 6. Task types in written tests	57
Table 7. Types of written assignments	58
Table 8. Task types in oral tests	59
Table 9. Types of oral presentations	60
Table 10. Portfolio assessment	61
Table 11. Methods used for assessing listening skills	62
Table 12. Methods used for assessing oral skills	63
Table 13. Methods used for assessing reading skills	64
Table 14. Methods used for assessing writing skills	65
Table 15. Methods used for assessing vocabulary	66
Table 16. Methods used for assessing grammar	67
Table 17. Methods used for assessing spelling	68
Table 18. Formative assessment	69
Table 19. Presentation of assessment results	70

Abstract

This study examines assessment methods in English in Icelandic primary and secondary schools along with how assessment results are presented to students and parents. The study views which methods are used for summative assessment and whether there is variety in methods used for the various language skills. The study also examines whether formative assessment is practiced and in what form, and whether there is a difference in assessment methods between age levels and/or in relation to teachers' education level and English teaching experience. The subjects of the study were English teachers of 11 secondary schools and 9 primary schools from all around the country. Data was collected through questionnaires via e-mail, and answers came from 24 secondary level teachers and 20 primary level teachers. Data was analysed statistically by the use of tables as well as verbal descriptions. The results show that written tests are the leading assessment method at both levels, usually counting for the majority of the final grade. However, some variety exists as all the participating teachers use other assessment methods as well. Listening and speaking skills are omitted in the assessment practices by some teachers, and formative assessment is currently not a high priority in the teachers' choice of assessment practices. Assessment results are presented with a grade in all cases, however, the majority of primary level teachers, and 41% of secondary level teachers, also give oral and/or written comments. There is no obvious correlation between teachers' education and experience, on the one hand, and choice of assessment methods on the other, however, an interesting difference appeared in how teachers' education and experience vary between age levels.

1 Introduction

Assessment is an important aspect of education. Schools work according to curricula and those curricula set goals towards which teaching is aimed. In order to find out whether students' progress meets expectations as laid out in the curricula, students, parents, and schools need some kind of assessment.

As a teacher with some twenty years of experience, I am well aware of the importance of assessment and how greatly it can affect those involved. In 2005, I read an interesting article by Auður Torfadóttir where the methods used for assessing foreign language learning within the Icelandic primary school system were discussed. This article touched me deeply, because through the years I had been struggling to find justifiable assessment methods that also were appropriate given the external framework and demands the school places on my work. Since reading the article I have been even more preoccupied with finding appropriate methods for assessing learning, and matching teaching approaches with learning goals and assessment.

Within education, the general consensus is that there has to be a relationship between teaching approaches and the assessment methods used in order for the assessment to be valid. Auður Torfadóttir (2005) claimed that it was her perception that assessment methods in foreign languages in Iceland are not in accordance with the teaching approaches advocated in the National Curriculum Guidelines for the Primary Level. This has to be a matter of some concern for the educational system on the whole. However, she noted that there were no studies to support this perception. Thus, a study that gives a description of the assessment methods used in today's schools in Iceland, has to be of interest for everybody involved, whether students, parents or teachers, theorists or educational authorities. Such a study could show whether these worries are justifiable or not, and whether there is a need for re-

evaluation of assessment within the Icelandic school system, at least where foreign languages are concerned.

Despite the acknowledged importance of assessment, there have not been conducted many formal studies of how it is performed in Icelandic schools, at least not when it comes to foreign languages. Erna Ingibjörg Pálsdóttir (2006) conducted a study of educational assessment in general in Icelandic primary schools, which clearly shows that there is a need for further development in the practice of assessment in Iceland (p. 119), thus drawing the attention to the need of further research within the field. Additionally, in describing a very recent study of methods used in assessing English learning in the 8th grade in Icelandic primary schools, Lilja Jóhannsdóttir (2008) argued that it was of interest to study assessment methods in Icelandic secondary schools and make a comparison between school levels (p. 58). Also, Guðfinna Gunnarsdóttir (2005), in studying teachers' grades and standardized test results in the 10th grade in primary schools in Southern Iceland, came to the conclusion that it would be of interest for the educational system to find out whether there is actually a difference between assessment methods at different age levels (p. 65). Hence, there are good arguments for the need of a study like the one reported here.

1.1 The Purpose of the Study

The purpose of this study was to describe the methods used for assessing English learning at all levels in Icelandic primary and secondary schools. Also, to examine the variety in assessment methods as advocated in the relevant Curriculum Guidelines, including the use of both formal and informal assessment, self-assessment as well as teacher assessment, and whether varied assignments are used to assess all language skills. Additionally, to view whether there is a difference in methodology depending on students' age levels, and whether teachers' education and/or experience has a noticeable effect on choice of methods. Finally, to

show how assessment results are presented to students and parents. Thus, the research questions are as follows:

- What methods are used for assessing English learning in Icelandic primary and secondary schools?
- Do the assessment methods used cover listening, speaking, reading, and writing skills according to learning goals?
- Is there variety in methods used to assess the different skills?
- Is formative assessment practiced?
- Is there a noticeable difference in assessment methods between age levels?
- Is there a correlation between methods chosen and teachers' education and/or experience?
- How are results of the assessment presented to students and, where appropriate, parents?

In order to answer these questions, questionnaires were sent to English teachers of 4 different grade levels at the Icelandic primary level, and 4 different course units at the Icelandic secondary level. The questionnaires asked about teachers' assessment practices and how the results of assessment are reported to students and parents, as well as basic information on their education and English teaching experience (see Appendix). Twelve schools from each educational level from all over the country were contacted in the first round, but because only 6 responded from the primary level, 8 additional primary schools were contacted. Teachers from 11 secondary schools, and 9 primary schools agreed to participate.

1.2 Organization of the Thesis

This thesis is divided into six chapters. Chapter 1 is the introduction, where the motivation for the study is explained, the purpose of the study is discussed and research

questions presented. Chapter 2 provides the theoretical foundation for the study through discussion of learning theories, different purposes and methods for assessment, possible impact of assessment, previous research of classroom-based assessment, and current policy in Icelandic National Curricula for primary and secondary schools. In chapter 3 the subject choice for this study is explained and data collection described. Chapter 4 is an account of the results of the study. Results are presented in tables to explain the statistical information from the questionnaires. In chapter 5 the results are discussed with references to the theoretical foundations presented in chapter 2. Chapter 6 provides conclusions of the study with some suggestions for further studies.

2 Literature Review

In this chapter the theoretical foundation for the study is reviewed. First, some definitions of terms used in discussing assessment are presented, and the connections between assessment and learning theories examined. Next, classroom-based assessment, or teacher assessment, is explained. Then different purposes of assessment are accounted for, methods used for assessing described, and current emphasis in assessment viewed. Following that, particular concerns regarding language assessment are considered, and possible effects on teaching and learning discussed. Then, the sections on assessment of the Icelandic National Curriculum Guidelines for both primary and secondary schools are examined. Finally, some previous studies of classroom-based assessment, of both international and Icelandic origin, are discussed.

2.1 Definitions of Terms

Since assessment has many faces, people use various terms in discussing it. It is thus important to define the main terms and constructs used in this study as thoroughly as possible.

2.1.1 Assessment and Measurement

Linn and Miller (2005) emphasized the significance of distinguishing between assessment, and measurement. They explained the term ‘assessment’ to be “a general term that includes the full range of procedures used to gain information about student learning ... and the formation of value judgments concerning learning progress” (p. 26). On the other hand, they claimed measurement to be “the process of obtaining a numerical description of the degree to which an individual possesses a particular characteristic” (p. 26). Bachman (1990) emphasized the quantifying role of measurement, and the importance of understanding that when measuring something, it is about abilities but not the persons themselves.

Furthermore, he claimed that measurement has to be performed according to rules so that it will not result in a random outcome (pp. 18 – 20). While measurement only is about quantitative descriptions, assessment can cover both quantitative and qualitative factors (Linn & Miller, 2005, p. 27).

Assessment can be formal and informal. Formal assessment is “assessment that is planned and carried out following formal procedures” (McKay, 2006, p. 20), and can be either external or classroom-based, whereas informal assessment refers to assessment which is performed “during the course of the teaching and learning process” (McKay, 2006, p. 20) and is solely classroom-based.

2.1.2 Classification

In 1972, Airasian and Madaus identified a system that classifies assessment according to its functional role, and divides it into four categories: placement assessment, which is meant to show students’ abilities at the beginning of course; diagnostic assessment, which is considerably specialized and used to diagnose any learning difficulties that might be troubling students; formative assessment, which is supposed to enhance learning progress during the course; and summative assessment, which is meant to assess students’ final achievement at the end of course (Linn & Miller, 2005, pp. 35 – 37).

A similar classification was made in 1988 by the Task Group on Assessment and Testing, where the categories are formative, diagnostic, summative, and evaluative and the first three are defined in much the same way as above. The fourth one, evaluative assessment, is used to report on the performance of some entities of educational systems such as a class or a school, and is based on summative assessment (Harlen, 2006a, p. 104). However, as Weeden, Winter and Broadfoot (2002) so eloquently put it, “it is not the assessment itself that is diagnostic, formative, summative or evaluative but the way that the information collected is

used” (p. 19). While these systems are still being used as classifications, new phrases have been invented to describe the formative and summative parts and those will be used interchangeably with the old ones from now on.

2.1.3 Summative Assessment

‘Assessment of learning’ (summative assessment) has been a part of education systems for centuries, and, in short, it is used “to summarize what has been learned” (Harlen, 2006a, p. 103). It seems that this kind of assessment is still the principal manner of assessment in schools (Earl, 2003, p. 22; Weeden, Winter & Broadfoot, 2002, p. 29). As explained by Linn and Miller (2005), assessment of learning is usually done at the end of instruction and the main purpose “is grading or the certification of student achievement” (p. 37). It is supposed to provide information about the students that can be used to inform those it concerns about the progress made, like the students themselves and their parents, or a new teacher or a school. Also, it is used to certify, in a formal way, the students’ achievement (Clarke, 2008, p. 13).

According to Earl (2003), the results of assessment of learning are usually given as number or letter grades, and often in a way that compares the students to others within a given group instead of being feedback that gives advice for further learning (p. 22 – 23). However, Gronlund (2003) has pointed out that assessment of learning ought to include some feedback to students (pp. 8 – 9), and Clarke (2008) has reported that the Assessment Reform Group, set up by the British Educational Research Association in 1989, emphasized in 2006 the need for assessment *of* learning to be in agreement with the process of assessment *for* learning (p. 12).

As explained by Linn and Miller (2005), various techniques, such as tests, performance-based assessment or portfolio assessment can be used to gather information used in assessment of learning, but the learning goals should always be kept in mind when deciding

which technique to choose (pp. 36 – 37). Since summative results can affect whether a student can continue at the next level of schooling it is of great importance that the assessment procedures are as valid and reliable as possible (McKay, 2006, p. 23).

2.1.4 Formative Assessment

‘Assessment for learning’ (formative assessment) has been defined by the Assessment Reform Group as “the process of seeking and interpreting evidence for use by learners and their teachers, to identify where the learners are in their learning, where they need to go and how best to get there” (as cited in Gardner, 2006, p. 2). Rea-Dickins (2007) has argued that where there is teaching which encourages student involvement with varied possibilities of language practice, for example through cooperation and self-assessment, and which involves different kinds of feedback to serve the students’ needs, there most likely is “good formative assessment practice” (p. 517).

As explained by Earl (2003), “assessment for learning happens in the middle of learning” (p. 24) and teachers play a very important role in it. They gather all sorts of information on the students’ learning and use various means such as “observation, worksheets, questioning in class, student-teacher conferences, or whatever mechanism is likely to give them information that will be useful for their planning and teaching” (Earl, 2003, p. 24). Assessment for learning is not meant to compare students, rather to pinpoint everybody’s strengths and weaknesses in order to contribute to their individual learning (Earl, 2003, p. 24).

For assessment for learning to work, there are a few strategies that have to be considered. These include making sure that the classroom culture is one where: ability is looked at as something that can be advanced; where effective dialogues are used and valuable questions are asked; where it is made sure that students understand learning objectives and are involved in planning; where students are actively involved in the assessment procedures

through self- and peer-assessment; and where there are ample opportunities for review and feedback as well as acting on the feedback (Clarke, 2008, p. 11).

Assessment for learning can be carefully planned, but can also be what has been called ‘on-the-run assessment’ which involves the informal observation and instantaneous feedback given by teachers in the classroom (McKay, 2006, p. 22). It is mainly used for pedagogic purposes, however, it also happens that teachers use the results to “devise a summative report” (McKay, 2006, p. 22) which can lead formative assessment to have high-stakes effects.

Recently, theorists have started to use the term ‘assessment as learning’ in addition to ‘assessment for learning’ to underline “the ultimate goal, where students are their own best assessors” (Earl, 2003, p. 25). Together, these two terms include everything that formative assessment is about and reach even further. As explained in a document developed by the Manitoba Education, Citizenship and Youth (2006), assessment as learning “emphasizes assessment as a process of metacognition” (p. 41). Also, it is made clear that assessment as learning is about getting the students to focus on, and control, their own learning (pp. 41 – 42).

2.1.5 Validity and Reliability

Regardless of which type of assessment is used, theorists agree that there are two crucial factors that have to be taken into account in all assessment. They are validity and reliability. Black and Wiliam (2006b) claim it to be essential that users of assessment results, whoever they may be, can rely on the results, and validity and reliability are the grounding for doing so (p. 119). Gronlund (2003) defined validity as the factor which determines the quality of the inferences and uses that are made from assessment results of any kind. He emphasized the fact that the important thing is the inference, not the assessment results themselves (p. 24). The likelihood that valid inferences are drawn depends, among other things, on adequate

sampling of the domain in question, proper selection and directions of tasks, sufficient time given, and adequate scoring criterion (Gronlund, 2003, pp. 203 – 205).

The second important factor, reliability, refers to “how consistent test scores or other assessment results are from one measurement to another” (Linn & Miller, 2005, p. 104). For any assessment results to be trustworthy there has to be some consistency between tests taken at different times or between various samples of the performance domain in question. However, one should not expect there to be perfect consistency since there can always be factors such as varied motivation, difference in test questions, personal biases in judgment of performance, or just pure luck that may bring errors into the picture (Linn & Miller, 2005, pp. 104 – 105). Nevertheless, it is really important for assessment that these errors be kept as slight as possible in order to make results as reliable as possible (Gronlund, 2003, p. 25).

Linn and Miller (2005) noted that one must always remember that reliability is a statistical notion that has to do with the results of assessment, and that reliable measure is not always valid (p. 131). And Gronlund (2003), while stating the importance of reliability, claimed that even though reliability is a necessary factor in acquiring valid inferences, it is not the only thing needed since it can also be invalid like when based on a consistent measurement of something completely different from what it is meant to assess (p. 25). To make sure that assessment has validity and reliability, intended learning outcomes must be defined beforehand (Gronlund, 2003, p. 25).

2.2 Learning Theories and Assessment

Through the years, assessment approaches have mainly relied on dominant learning theories at each time. According to Shepard (2000), the predominant theory of assessment in the early 1900s was based on the beliefs of the social efficiency movement. The movement’s aim was to minimize waste within the school system, hence leading to a curriculum where

academic subjects were eliminated except for those that were believed to have relevance for college studies (pp. 4 – 5).

In the 1920s, the advance of tests relied on behavioural psychology and learning was seen as being first and foremost hierarchical where basic skills have to be mastered before any higher-level skills can be learnt (Linn & Miller, 2005, p. 50). Furthermore, as claimed by James (2006), behaviourism stayed the principal theory of learning through the 1970s and can even be said to be affecting schooling and assessment today (p. 54). Additionally, James (2006) explained that behaviourism sees learning “as the conditioned response to external stimuli” (p. 54), where the use of reward systems is believed to help achieving set goals through practicing bit by bit of a complex whole and always building upon earlier practices (p. 54). Also, behaviourist theories do not see human consciousness as having anything to do with how learning happens. As for assessment, these theories have created a way that includes “unseen timed tests with items taken from progressive levels in a skill hierarchy” (pp. 54 – 55) where the answers are usually only correct or incorrect.

According to James (2006), new theories started to emerge in the 1960s as a response to behaviourism. Those theories are called cognitive or constructivist theories and according to them the human mind actually plays a role in learning; learning is dependent on mental construction and active participation of the learner. Also, as James (2006) wrote, “a particular focus is on how people construct meaning and make sense of the world through organizing structures, concepts and principles in schema” (p. 55). Furthermore, understanding and problem solving are considered principal factors, as well as self-monitoring and self-regulation, and these theories consider prior learning to be of great importance for new learning. James (2006) explained that these theories have gained much attention in recent years because of how they affect teaching and assessment. Formative assessment has its

foundation in these theories where the goals of learning are the most important aspect (pp. 55 – 56).

The socio-cultural theory, a very influential theory of learning and assessment, is originated in Vygotsky's thinking within psychology but also has its roots in social theory, sociology and anthropology. This theory views learning as taking place through interaction. The essential notion is that various cultural artefacts, such as language, affect the mediated activity that learning actually is (James, 2006, pp. 56 – 57). Furthermore, James (2006) noted that the socio-cultural theory emphasizes the importance of a learning environment where students can expand their knowledge through scaffolding made by their teachers or peers within what is called their zone of proximal development, meaning when learners get to deal with tasks they would not be able to manage on their own. However, she claimed that assessment has not yet been thoroughly worked out within this theory. The focus has been on assessing how well the students are able to use the means “available to them to formulate problems, work productively, and evaluate their efforts” (p. 58) and portfolios have been seen as effective ways for the assessment even though there are problems with the actual grading.

Shepard (2000) has suggested a social-constructivist framework for assessment based on the cognitive, constructivist, and socio-cultural theories in order to make classroom assessment more effective, and further learning processes (p. 6). She claimed that in order to follow this social-constructivist framework, it is essential that classroom assessment practices change in two ways. Form and content have to base on “important thinking and problem solving skills in each of the disciplines” (p. 7), and there has to be a change in how teachers and students view and use assessment. Moreover, she stated that in order for the latter to happen, teachers have to be helped getting rid of the negative effects that external, high-stakes testing has created (pp. 7 – 9) and will be further discussed in chapter 2.7.

2.3 Classroom-based Assessment

The form of assessment which interests the study in question here is classroom-based assessment. Classroom-based assessment is a term used for assessment that is performed within a classroom as a part of the teaching process, also referred to as teacher-based assessment (Hamp-Lyons, 2007, pp. 488 – 489). In distinguishing between classroom-based and external assessment, Hamp-Lyons (2007) has stated that:

The contexts and needs of classrooms and teachers are not the same as those of large scale testing. The large scale needs to discriminate, to separate, to categorize and label. It seeks the general, the common, the group identifier, the scaleable, the replicable, the predictable, the consistent, and the characteristic. The teacher, the classroom, seeks the special, the individual, the changing, the changeable, the surprising, the subtle, the textured, and the unique. Neither is better but they are different. (p. 487).

As with any other form of assessment, there are various purposes for classroom-based assessment “ranging from meeting the bureaucratic demands placed on teachers for data on learner achievement levels to assessment that has a primarily supportive function in the formative assessment of language learners and is firmly embedded within routine instructional contexts” (Rea-Dickins, 2007, p. 508), and these varied purposes lead to different perceptions of assessment. And, as explained by McKay (2006), classroom-based assessment, which often is looked upon as solely formative, can also include summative assessment, be it carried out by the teacher in the classroom, and it can actually be described as a continuum with formative and summative assessments on each end (p. 68).

Classroom-based assessment is a complex notion, which needs to be given a lot of thought and practice. Rea-Dickins (2007) has argued that teachers need to consider some potential dangers when applying classroom-based assessment. One pitfall to avoid is the use

of “a series of summative mini-achievement tests” (p. 513). Another one is failing to test all the language skills; speaking and listening, for example, seem to be left un-assessed in many cases (p. 513). And when it comes to assessing young language learners’ progress, it is vital to keep in mind the learners’ age, their cognitive and linguistic skills, their interests, and the expected attention span, when planning the classroom assessment (Zangl, 2000, p. 251). Also, for teachers to be able to carry out classroom-based assessment with some reliability, they have “to have some understanding of the process of L2 acquisition” (Zangl, 2000, p. 256).

Recent studies into classroom-based assessment in foreign languages show that there can be some inconsistencies, leading to lack of validity and reliability, which is of some concern since high-stakes decisions are sometimes based on classroom assessment (McKay, 2006, pp. 68 – 69). However, it seems that validity and reliability have to be considered in a somewhat different way when it comes to classroom-based assessment. McKay (2006) has claimed “that the main yardstick for validity in classroom formative assessment is the success of the learning that takes place” (p. 358). She also maintained that by gathering a sufficient amount of information in classroom formative assessment, reliability would be secured (p. 358), thus agreeing with Hamp-Lyons (2007) that classroom-based assessment should be considered as reliable and as trustworthy as external testing (p. 501). In explaining that “some of the fundamental assumptions of reliability theory are not appropriate for classroom assessment” (Hamp-Lyons, 2007, p. 501), Hamp-Lyons used as evidence the fact that teachers naturally want their students to develop their knowledge and skills from week to week, thus consistent measures can not be expected across time.

2.4 Purposes and Methods

“Assessment works best when its purpose is clear, and when it is carefully designed to fit that purpose” (Manitoba Education, 2006, p.13). However, Earl (2003) claimed that

assessment is a complicated entity with various purposes, all of which need to be understood and appreciated for their own merits. In explaining this further, she noted that in 1998 Black claimed the main purposes of assessment to be threefold: to enhance learning, to report on and certify student achievement, and public accountability (p. 12). Also, Earl (2003) emphasized that given these different assessment purposes, varied approaches for assessment are necessary (pp. 12 – 13). It is essential that teachers know and understand each purpose, and are aware of when and why they are using which so that they make a valid choice of approaches (Manitoba Education, 2006, p. 14). Thus, it is obvious that the purpose of assessment should always be the centre of attention and as argued by Black and Wiliam (2006a), “assessment in education must, first and foremost, serve the purpose of supporting learning” (p. 9).

Through the years, various methods and approaches for assessing learning have been used and it is necessary to understand the benefits and possible defects of each, as well as be aware of for what purposes each one fits the best. Stiggins (2004) wrote that “if teachers assess accurately and use the results effectively, then students prosper. If they do it poorly, student learning suffers” (A Legacy of Mistaken Beliefs section, para. 23). Thus, choosing the appropriate type of assessment is of great relevance. Below, the following assessment methods will be discussed: tests, performance-based assessment, portfolio assessment, self-assessment and peer-assessment.

2.4.1 Tests

Testing has long been the most common assessment method. Linn and Miller (2005) described a test as an instrument used for assessing, usually made of some questions and put in front of students under some particular circumstances that can be said to be comparable for all participants (p. 26). Furthermore, Bachman (1990) stated that tests distinguish themselves

from other measuring methods used in language assessment since they are meant to gather *specific* samples of language use from which inferences can be made of the test-taker's language abilities, hence justifies the use of language tests (pp.20 – 21). While Bachman and Palmer (1996) believe that language testing is a valid way to interpret the language abilities of individuals, they also claim that teachers have to be aware of the importance of how tests are made (p. 8). In addition they put forward what they call their philosophy of language testing:

- 1 Relate language testing to language teaching and language use.
- 2 Design your tests so as to encourage and enable test takers to perform at their highest level of ability.
- 3 Build considerations of fairness into test design.
- 4 Humanize the testing process: seek ways in which to involve test takers more directly into the testing process; treat test takers as responsible individuals; provide them with as complete information about the entire testing procedure as possible.
- 5 Demand accountability for test use; hold yourself, as well as any others who use your test, accountable for the way your test is used.
- 6 Recognize that decisions based on test scores are fraught with dilemmas, and that there are no universal answers to these. (p. 13).

It is necessary to remember that a test can only contain a limited amount of what has been covered with the instruction (Earl, 2003, pp. 22 – 23) and the results only show what the student can do at the exact time the test is completed (Weeden, Winter, & Broadfoot, 2002, p. 29). Thus, it is important to choose samples for each test with great care, and always take the circumstances into account when the results are interpreted.

2.4.2 Performance-based Assessment

Another method commonly used in assessment today, not the least language assessment, is the so-called performance-based assessment. According to Linn and Miller (2005), a movement towards new approaches to assessment developed in the 1990s, and these approaches have been “referred to as alternative assessment, authentic assessment, direct assessment, or performance-based assessment” (p. 7). Also, they explained that ‘alternative’ emphasized the importance of changing assessment so that it focused on something other than multiple-choice tests, with ‘authentic’ came the demand of having tasks that could be accounted relevant outside the classroom world, and ‘performance’ highlighted the importance of “the actual doing of a task” (p. 7).

McKay (2006) points out that the use of “selected-response items is avoided” (p. 98) in performance-based assessment. The focus is on tasks that enable the learners to use the language as they would in real situations, where both grammar and vocabulary knowledge are assessed as parts of the communicative use of the language (p. 99). Scoring rubrics are recognized as an appropriate way to assess performance since the tasks used are not likely to have only correct or incorrect answers; however, one can not expect any assessment criteria to be so precise that no interpretation by teachers will be needed (McKay, 2006, pp. 266 – 268).

Even though performance-based assessment is relatively new within the field of general educational assessment, it is something that language assessors are quite familiar with and has been used in language assessment for more than a century (Bachman, 2002a, p. 5).

2.4.3 Portfolio Assessment

Portfolio assessment is currently a very popular alternative assessment method where instruction and assessment are integrated (Hamp-Lyons, 2007, p. 493). According to McKay (2006), portfolios have, for many years now, been considered important in assessment in

primary schools and are “widely advocated” (p. 159). In some schools, portfolios are the sole method used for assessment while in others they are used along with other methods (Gronlund, 2003, p. 157). In explaining portfolios, which are collections of various student work accumulated over the study period, McKay (2006) emphasized that for using them for assessment, samples have to be “systematically collected” (p. 159) and the students should be involved in the selection of content. A significant aspect of the portfolio assessment is student self-reflection, which enables students to understand where they are in their learning, by which means they have reached that position, and where to go from there (McKay, 2006, p. 159) which rhymes very well with formative assessment.

One of the reasons for why portfolios are considered to be advantageous for assessment is the variety of evidence they can show to judge student performance:

1. Learning progress over time.
2. Student’s current best work.
3. Comparison of best work to past work.
4. Development of self-assessment skills.
5. Development of reflective learning.
6. Individual’s level and pace of work.
7. Clear evidence of learning to parents and others.
8. The amount of teacher-student collaboration involved. (Gronlund, 2003, p. 158).

However, applying portfolios is quite time consuming and demands substantial teacher-student conferencing if it is supposed to be a useful tool (Gronlund, 2003, p. 159).

As explained by McKay (2006), portfolios can be of different types, with different purposes. First there is the process portfolio, which is intended to collect information to show students’ progress from day to day in the view of short-term objectives. Then there is the archival portfolio, also called showcase portfolio, in which products are selected a few times

over the study period in order to show the students' abilities, thus present "the basis for summative assessment" (p. 160). Finally, McKay (2006) discusses the aggregated portfolio, which "is concerned with evidence for accountability and for evaluation of the programme" (p. 161) based on samples from each of the students in the study group.

As for evaluating the portfolio, Gronlund (2003) stated that the criteria must be decided beforehand and then "converted to rating scales or other scoring rubrics that can be used in self-assessment, peer assessment, and teacher assessment" (p. 163). He also emphasized that both "the portfolio structure and the students' overall performance" (p. 163) need to be evaluated.

2.4.4 Self-assessment

Self-assessment is one of the strategies which formative assessment relies on, and in fact Black, Harrison, Lee, Marshall, and Wiliam (2003) have asserted that it is of fundamental value for the development of formative assessment that self-assessment be practiced by students (p. 18). According to McKay (2006), self-assessment is a way to persuade students to focus on their own learning in order to better understand the process and to accept the responsibility for it, and self-assessment sheets of some kind are often kept as part of students' portfolio (pp. 165 – 166). One of the things that make self-assessment important is that it usually views the students' learning in comparison with *their* earlier achievements instead of other students' (Weeden, Winter, & Broadfoot, 2002, p. 77).

Linn and Miller (2005) have claimed that it can be beneficial for students' learning to have them rate themselves and then, through a teacher-student conference, "compare the ratings with those of the teacher" (p. 275). Furthermore, they maintained that students should be involved in creating criteria and/or instruments to use in the rating process. Such procedure is likely to give the students more insight in their own studies, as well as help them develop

better self-assessment skills (p. 275). For any learning to take place, the learners have to understand the learning goals as well as how best to work to reach them, thus self-assessment enhances learning (Black & Wiliam, 2006a, p. 15). “Teaching, learning and assessment are always closely linked, but in using self-assessment strategies a teacher further blurs the dividing line, as a pupil’s learning activities will involve aspects of assessment of those activities” (Weeden, Winter, & Broadfoot, 2002, p. 79).

2.4.5 Peer-assessment

Peer-assessment is also one of the methods used in formative assessment, and, according to Black et al. (2003), it is “uniquely valuable” (p. 50) and “may even be a prior requirement for self-assessment” (p. 50). It is suggested by many researchers that peer collaboration in classroom activities are beneficial to student learning, and this they sustain with evidence from various research. Additionally, evidence shows beneficial effects of peer-assessment; however, it also shows that in order for peer-assessment to be of use, students have to know how to execute it (Saito, 2008, pp. 553 – 554). So, as with self-assessment, students need to develop assessment skills; they have to learn how to use criteria and how to treat their peers with respect (McKay, 2006, p. 166).

Like self-assessment, peer-assessment stimulates autonomy through reflection on learning and objectives, as well as viewing problems and possible solutions to those (James & Pedder, 2006, p. 28). Furthermore, students will benefit in more ways from this practice; they will, among other things, “increase their language awareness and ability to talk about language” (McKay, 2006, p. 166), and their personal and social development will be positively affected by it (Weeden, Winter, & Broadfoot, 2002, p. 89). However, despite the benefits which undeniably are related to peer-assessment, McKay (2006) emphasizes that

teachers need to be aware of the limitations of children's ability to "give an abstract, independent assessment of a performance" (p. 166).

2.5 Putting Formative Assessment into Practice

Current literature on assessment mostly focuses on the importance of assessment *for* learning, and as claimed by Stiggins (2007), schools in today's world see their job to be about enhancing the achievement of all their students instead of yesteryear's ranking tradition (22). Research evidence shows that formative assessment can improve student achievement if its practice is supported and developed with the sole intention to further learning (Black & Wiliam, 2006a, p. 9), thus, this increasing interest in implementing formative assessment is understandable.

Nevertheless, according to Black et al. (2003), surveys of teacher practice have shown that efficient formative assessment is actually not commonly used in classrooms today (p. 2). Thus, they started a project with a group of teachers in British schools where the main goal was to develop effective assessment for learning within the classroom work that could become a conventional practice for teachers (p. 17). The outcome was a development of approaches within four areas, questioning; feedback, with emphasis on marking; self-assessment, along with peer-assessment; and the formative use of summative tests. (p. 31).

As for questioning, research has shown that the average time teachers wait after having asked a question is less than a second, whereas extended wait time results in longer and more confident answers, as well as more answers on the whole. Transcripts of lessons in this project show a remarkable gain in students' thinking and participating when the teachers used open questions and wait time was extended (Black et al., 2003, pp. 32 – 39). The second aspect, feedback, also developed in positive ways. If it is to be effective, feedback has to include proposals for what to do to improve one's work, while general statements like 'good

job' do not do the trick (Brookhart, 2007, p. 56). Black et al. (2003) noted that most of the comments made by participating teachers at the beginning of their project did not give any advice as to which steps to take for improvement. However, they noticed a considerable change through the project where the teachers' comments progressed and became more in the line of effective feedback, leading to enhanced learning (pp. 44 – 45). Also, the feedback led to more understanding of learning goals and ways to reach them, thus initiated self-assessment (pp. 49 – 50).

The teachers involved in the project believed it was necessary to find a way to use summative tests formatively, since they felt the tests were a necessary part of their assessment practices. They used formative strategies in preparing for tests so the students would focus their revision on areas of potential weakness; they used test results in a formative way, where they took areas which did not get successful answers and concentrated on finding better ways for students to learn those things; and they used peer marking of tests where the students first had to create a scheme to use for the marking, so they needed to think about relevant criteria. From the findings of this project, one can see that summative tests can be used in positive ways to enhance student learning (Black et al., 2003, pp. 53 – 56).

One of the results of this project was more autonomous students who, through increased involvement in the study process, started to understand when they were actually learning and how to get there, and that, according to Black et al. (2003), can be looked at as “one of the most important benefits of formative assessment” (p. 67). Furthermore, they argued that since their findings are coinciding with extensive research material within education and psychology, it is likely that they can be generalized (p. 77).

As noted before, formative assessment is not commonly practiced in schools, regardless of all the research showing its benefits. A likely explanation is teachers' workload. As explained by Black et al. (2003), the implementation of formative assessment cannot

happen without changes in the ways that teachers view their roles and what they consider as being good teaching (p. 80). These changes are demanding since they require the teachers to take on whatever role needed to enhance student learning, and those can be of various kinds (p. 99). Hamp-Lyons (2007) has, for example, stated that teachers need considerable assistance in developing the skills of using effective feedback in language teaching (p. 492). Also, Wiliam (2007) has claimed that “teachers need substantial support and guidance to integrate formative assessment into their practice” (p. 41), it is not enough to inform them about the positive outcome of research regarding assessment for learning; their actual actions have to be the main focus. Wiliam (2007) has suggested an approach that has been found to be successful and involves learning communities for teachers that meet regularly, and develop plans for their changes (pp. 38 – 40).

2.6 Concerns Regarding Language Assessment

“Assessing language is not like assessing maths or geography or physics. Language is a construct that is hard to define, although we are surrounded by it and immersed in it, and we know it when we see/hear it” (Hamp-Lyons, 2007, p. 488). In accordance with this, Bachman (1990) has stated that “the challenges facing language testers are immense” (p. 357), which can be viewed as a frightening statement. Still, he claimed that it is actually exciting to face these challenges since it will hopefully result in better understanding of language abilities and language use, as well as more useful language tests (p. 357).

As with any other assessment, the principal consideration in language assessment is its purpose, and the main purpose of language testing is information gathering in order to be able to interpret the test taker’s language abilities (Bachman & Palmer, 1996, p. 23). However, as noted by Bachman (2002a), in language testing one has to deal with the conflict that language is the object of the assessment as well as the instrument used (p. 5) which undeniably makes

things more complicated. And since “language is difficult to characterize, ... proficiency in the language will be difficult to measure” (Hamp-Lyons, 2007, p. 488).

Bachman (1990) stated that in developing language tests one has to consider all the factors that affect the test taker’s performance, and that in order for tests to actually assess what they are intended to they have to rely on clear definitions of the two essential components that affect the results: the language abilities and the test methods (p. 81).

Bachman and Palmer (1996) asserted that their approach to language test development is based on “the need for a correspondence between language test performance and language use” and “a clear and explicit definition of the qualities of test usefulness” (p. 9). Additionally, they proposed a model for finding how useful a test is for its intended purpose. The model consists of six important test qualities: “reliability, construct validity, authenticity, interactiveness, impact, and practicality” (p. 17). Here, Bachman and Palmer (1996) consider the importance of authenticity to be about the possibilities of generalization, and in language testing one of the main purposes is to be able to generalize the interpretations about language ability to a target language use domain (pp. 23 – 24).

According to Alderson and Banerjee (2002), this model, called the Bachman model, has been very influential in language testing and is, by many, considered to be very useful in test construction (p. 80). Nevertheless, the Bachman model has also been challenged. As explained by Alderson and Banerjee (2002), McNamara claimed that this model is mainly based on psychological theories of language use, thus ignoring the influential social aspects which he believes have to be considered (p. 81). Also, it is not clear how the different components of the model interrelate and should be evaluated (Alderson & Banerjee, 2002, p. 80).

Bachman and Cohen (1998) noted that in the 1960s and 1970s language was mainly seen from a linguistic point of view and language testing focused on the linguistic aspects

such as grammar and vocabulary, whereas in the 1980s a more communicative view of language ability claimed its place as a result of a widening in the theoretical grounding for second language acquisition research (pp. 4 – 5). This resulted in a more complex construct for language assessment, since the sociolinguistic and pragmatic aspects of the language needed to be assessed as well as the grammatical ones. Thus, performance-based assessment became important. Bachman (2002a) has claimed that when it comes to performance-based assessment of language ability, “the topical content of assessment tasks and the knowledge schemata of test takers play much more important roles than on selected-response task types” (p. 16). Hence, it is important to base inferences of results on construct definitions of both language ability and topical knowledge.

According to Bachman (2002b), performance-based assessment usually measures complex abilities in a more complex way than “traditional constructed-response items” (p. 471). Furthermore, Bachman (2002a) discussed the difference between task- and construct-centred assessment methods, which both can be said to have certain disadvantages, and came to the conclusion that one should not exclude the other; both are needed for high-quality assessment (p. 16). Bachman (2002b) has argued that there is agreement among theorists that the most important issue in language performance assessment is to understand effects that assessment tasks have on test performance, since it is connected to how test results can be interpreted and used as well as for the validity of the interpretation (p. 468).

As for special considerations regarding assessment of particular skills in foreign languages, Alderson and Banerjee (2002) have discussed those. According to them, theorists are very interested in researching whether reading in a foreign language is actually about reading abilities or language abilities, and the answer is as yet unknown (pp. 83 – 84). Also, they claim assessment of listening, the least understood aspect of language testing, only to be possible through involving other skills, such as reading and writing, or speaking, as well (pp.

87, 89). Assessing speaking abilities in a second language became a matter of interest when people started to view language within a communicative approach. However, as noted by Alderson and Banerjee (2002), “speaking tests are particularly problematic from the point of view of reliability, validity, practicality and generalisability” (p. 92) and there is ongoing research interest in that area, including some cross disciplinary research concerning applied linguistics, discourse analysis and second language acquisition. Furthermore, assessment of speaking and writing abilities has been confronted with similar problems since the evaluation of writing started to include the discourse structure of the writing in addition to grammar and vocabulary (p. 95).

2.7 Impact of Assessment

It has been argued that “teachers can enhance or destroy students’ desires to learn more quickly and more permanently through their use of assessment than through any other tools at their disposal” (Harlen, 2006b, p. 62). Thus, assessment can be quite powerful, and it has been seen to have both positive and negative effects on those it is applied to (McKay, 2006, p. 18). McKay (2006) explained positive effects to be the results of effective assessment, where students benefit from the procedure, where it is secured as possible that what is supposed to be measured is measured without bias, and where teachers, and others involved, get valuable information about the process of teaching and learning (pp. 18 – 19).

However, not all assessment can be said to have positive effects. McKay (2006) has claimed that assessment can lead to undesirable power relationships, and favour some students at the cost of others, thus have negative effects (p. 19). Also, there are the so-called high-stakes assessment results. As explained by Harlen (2005), test results become high-stakes results when used to make important decisions that affect not only the individual student but the school system on the whole. Research has shown that it puts pressure on the

teachers that greatly influences the teaching and students' learning experiences, as well as resulting in very limited possibilities for learning. In addition, she stated that even though statistics show that the intense focus on passing tests has resulted in raised scores, there is research evidence showing that it has *not* resulted in raising students' achievement. This, as she explained, is due to a number of reasons. Among others is the fact that tests are meant to be objective, which, since everything has to be easily marked, leads to an elimination of the need to use problem solving and critical thinking skills in the assessment process (pp. 208 – 209).

Harlen and Deakin Crick (2003) noted that at the turn of the twenty-first century there was research evidence showing that external testing was holding formative assessment back and had negative effects on motivation for learning, especially for those students who were less successful at school than others (p. 170). Furthermore, they claimed that Kellaghan et al. had, in 1996, supplied evidence showing that when students were motivated by external tests they were more likely to have performance goals than learning goals, meaning that they were shallow learners who depended on rote learning (p. 171). Also, Harlen (2005) drew attention to other negative effects that high-stakes testing has, which have to do with motivation for learning. These are, for example, serious aspects such as provoking test anxiety and low self-esteem. Harlen's conclusion is that it is a real worry for education when negative influences on motivation for learning are disregarded, since lifelong learning is considered to be of great importance (p. 210).

2.8 Icelandic National Curriculum Guidelines

The current National Curriculum Guidelines for the Primary Level is from 2007 and is still in the process of being adopted by schools. The previous one, from 1999, is valid until 2010 and it varies how soon schools take the new one up (Menntamálaráðuneytið, n.d.b).

However, since there is no noticeable difference in the policy of assessment in the general section of these two editions, a choice was made to refer to the new one here. As for the National Curriculum Guidelines for Foreign Languages at the Primary Level, there is some change in the way that the new edition has more thorough explanations. Since this study examines the assessment methods used in the 4th grade, in which English is a compulsory subject according to the new edition but not the older one, the new one will be the foundation for discussion here. The current National Curriculum Guidelines for Foreign Languages at the Secondary Level date from 1999, whereas the general section is from 2004. Hence, these Curriculum Guidelines should already have been adopted by all schools (Menntamálaráðuneytið, n.d.c).

2.8.1 The Guidelines for the Primary Level

In line with the main emphasis advocated by assessment theorists today, the National Curriculum Guidelines for the Primary Level claim the main purpose of assessing learning to be to encourage students to do their utmost in their studies; thus enhance learning. Furthermore, it is said to be an important way to accumulate data that can be useful for the future organization of the studies. Assessment is considered to be an aspect of the education system that can not be set apart from the other areas; teaching and studying, hence continuous assessment is advocated (*Aðalnámskrá grunnskóla*, 2006, p. 16). It is emphasized that assessment methods must always be in line with the teaching approaches which hopefully are in line with the learning goals where all language skills are included. Assessment thus cannot build solely on formal methods. Moreover, the Curriculum Guidelines state that all assessment should be fair and honest so that all skills have to be considered in accordance with the instructional methods and emphases. Additionally, placement assessment is recommended at the beginning of a study period to give the opportunity for a purposeful

planning. Also, the importance of coaching students towards self-assessment is highlighted (*Aðalnámskrá grunnskóla*, 2006, p. 17).

In addition to what the Curriculum Guidelines have to say about assessment in general, the new National Curriculum Guidelines for Foreign Languages at the Primary Level emphasize the importance of continuous assessment since it is well suited for language assessment. Also, it is pointed out that a logbook can be of much help, and the European Language Portfolio is recommended in order to develop self-assessment skills. The use of rubrics is recommended so that students will get as thorough information as possible concerning their language abilities, both the strong and the weak sides, and examples are given of possible assessment tasks (*Aðalnámskrá grunnskóla*, 2007, pp. 13 – 15).

As for the assessment of English in particular, the Curriculum Guidelines state that formal assessment is not suitable for young learners. For the 4th grade, written comments, built on what the students *can* do, are recommended instead of grades (*Aðalnámskrá grunnskóla*, 2007, p. 17). In the 5th grade, the main emphasis is on listening and speaking and the assessment has to be in accordance with that. In the 6th and 7th grades, reading abilities need to be considered as well as listening and speaking, while written work should not be assessed according to highest standards even though the students are expected to have begun the developing of writing skills. Here, written feedback is also recommended in order to encourage and inform the students. In the 8th – 10th grades, all aspects of the language should be assessed according to the emphasis of the studies, and the assessment has to give the best possible information based on reliable sources. It is also underlined that since standardized tests can not cover all the goals set, it is important that the teacher assessment covers what is left, especially where oral abilities are concerned (*Aðalnámskrá grunnskóla*, 2007, p. 18).

Regarding the presentation of assessment results, the Curriculum Guidelines underline the significance of clear demonstration so students and parents can easily understand the

implications of the results. However, schools are at liberty to choose whether they use numbers, letters or words as long as the criteria used are known to everybody involved. Also, the Curriculum Guidelines highlight the importance of oral feedback along with written feedback, especially for the younger students (*Aðalnámskrá grunnskóla*, 2006, p. 18).

2.8.2 The Guidelines for the Secondary Level

At the secondary level, the main purpose of assessment in general is to find out to what extent students have met the goals set for a particular subject area and there is a special section regarding testing in particular (*Aðalnámskrá framhaldsskóla*, 2004, p. 33). The National Curriculum Guidelines for Foreign Languages at the Secondary Level recommend continuous assessment instead of only assigning a final test, and emphasize the importance of formative assessment. Furthermore, the importance of assessing all the aspects covered by the learning goals for each course is highlighted, and there is a reminder of balance in assessment and not making some aspects weigh too much while others get too little attention (*Aðalnámskrá framhaldsskóla*, 1999, p. 15). The Curriculum Guidelines emphasize the importance of practicing and assessing all language skills for all the course units discussed in this study apart from ENS 503. There, reading and listening skills are important elements but the emphasis is primarily on written work (*Aðalnámskrá framhaldsskóla*, 1999, pp. 29 – 39; *Aðalnámskrá framhaldsskóla*, 2000, pp. 11 – 13). Also, the Curriculum Guidelines call attention to the importance of varied assessment methods and examples are given of diverse possibilities for assessing each language skill (*Aðalnámskrá framhaldsskóla*, 1999, p. 21). Finally, the Curriculum Guidelines underline that in the light of globalization it is natural to have international standardized tests in mind while planning English studies at the secondary level, naming especially the Cambridge Examinations: First Certificate in English, Certificate

in Advanced English, and Certificate of Proficiency in English (*Aðalnámskrá framhaldsskóla*, 1999, p. 22).

Concerning the presentation of assessment results at the secondary level, the Curriculum Guidelines give orders for grades to be given in numbers on the scale of 1 to 10, and there are some requirements regarding the information that should be included on the final certificate. However, there are no particular requirements for the presentation of assessment results after each course unit, apart from those concerning the protection of personal identities (*Aðalnámskrá framhaldsskóla*, 2004, pp. 34 – 35).

2.9 A Review of Previous Research on Classroom-based Assessment

2.9.1 International Research

There is extensive international research within the field of educational assessment in general (Broadfoot & Black, 2004, p. 8), whereas, according to Cheng, Rogers, and Hu (2004), studies of assessment practices within foreign or second languages are more limited (p.360). In the light of the aims of my study, I have chosen to discuss a few of the recent ones that mainly focus on classroom-based assessment of the study of English as a foreign or second language.

Rea-Dickins and Gardner (2000) conducted a study in nine primary schools where they, among other things, investigated the assessment procedures implemented by teachers and what effect they had. They found that there was quite some variety in the assessment methods, as well as the use of the data gathered from assessments (p. 220). However, they also found that even though decisions were mostly based on multiple data gathered continuously through various sources, it did not necessarily lead to valid inferences about learners' language abilities since some of the data used was unreliable. Those were inaccurate

transcriptions of language samples which the teachers noted down in the midst of the learners' class discussions (pp. 234 – 236).

Cheng, Rogers, and Hu (2004) conducted a comparative survey in Canada, Hong Kong, and China where the purposes, methods, and procedures of assessment within the tertiary level were studied (p. 363). Their findings included a noticeable difference in methods for assessing reading between the countries, where in China it was more common to use multiple-choice questions and standardized reading tests, while in Canada and Hong Kong the emphasis was on student summaries of what was read and short answer questions. Also, the use of portfolios was more common in Canada than in Hong Kong and China. However, when it came to listening and speaking, there was not a great difference between the countries. Average class sizes in China are twice and three times larger than in Hong Kong and Canada respectively, which is interpreted as influencing assessment choices. Furthermore, teachers in China are found to have less experience in teaching and less training in assessment than those in Canada and Hong Kong, and that too is considered to have effects on assessment choices. The most frequent feedback given in all three settings is oral feedback, even though it is not as commonly used in China as in Canada and Hong Kong. The final reporting of assessment results was most commonly in the form of total scores and letter grades, accompanied with written comments in about half of the cases (pp. 378 – 379).

Barootchi and Keshavarz (2002) conducted a study of portfolio assessment in an Iranian high-school, with an experimental group of 30 students where portfolio assessment was practiced and a control group of other 30 students. The purpose was to find out “whether portfolio assessment contributes to EFL learners' achievement and their feeling of responsibility for monitoring their own progress” (p. 282). Also, to view whether there was “correlation between the scores of portfolio assessment and those of teacher-made tests” (p. 282). The correlation was measured with a correlation coefficient test and a reasonable

correlation was found (p. 284). The findings of the study also show that portfolio assessment can be of great value in enhancing learners' achievement since there was a considerable difference between the achievement scores of the two groups in favour of the experimental group, when measured with a t-test (p. 285). Additionally, this study indicates that students view portfolio assessment positively and it promotes their responsibility for their learning. Barootchi and Keshavarz (2002) conclude that only portfolio assessment can serve the purpose of making students conscious of their progress and responsibility, whereas a combined use of portfolio assessment and tests can serve the needs for both assessing specific language abilities and continuous student progress (p. 286).

Leung and Mohan (2004) have claimed that formative assessment is a promising way for assessing learning; however, they believe that the realization of formative assessment through teacher-student discourse in the classroom needs to be researched further. They conducted a case study in two classes in primary schools in England in order to study the interaction between teacher and students in formative teacher assessment within the classroom (pp. 335 – 336). They found that the teachers, while having the students discuss possible answers by opting possibilities and giving arguments for their acceptance or rejection of those, were actually helping the students develop academic discourse (p. 356). The teachers did not treat answers as final, rather encouraged students to discuss and rethink their answers, thus, used scaffolding to make the students reason about their decisions (p. 355).

Furthermore, Leung and Mohan (2004) developed a three-move analytical framework to use in classroom-based teacher assessment. The three moves are “offers, reasons and responses” (p. 357). This frame can clarify the form of the teachers' scaffolding as well as the participation of the students, and Leung and Mohan (2004) claimed that using this analytical approach “can shed light on issues such as teacher interpretation of assessment criteria, descriptors and, importantly, notions of language development and progress” (p. 357).

One of the debates that have been ongoing regarding methods used in assessing foreign languages is the question of whether to allow access to dictionaries during tests. The debate is affected by two different theories, the constructivist process-based approach and behaviourist knowledge-based approach. The former is considered to be more learner-centred since it is meant to enhance learning, as well as learner autonomy and motivation, whereas the latter is viewed as the traditional summative way of obtaining static proof of students' abilities. Hence, the constructivist approach is inclined to see the use of dictionaries as positive while the behaviourist approach sees it as negative (East, 2007, pp. 332 – 333).

Studies on this have not yielded unanimous results. East (2007) has reported on a study, conducted in 2003, that examined whether bilingual dictionary use in writing tests affects test scores (p. 332). This study included “17 – 18 year old students in New Zealand secondary schools” (p. 338) studying German. The study shows no significant difference to scores whether a dictionary was allowed or not (p. 346), and students' experience with a dictionary did not seem to matter either (p. 344). These findings are in contrast to those of Hurman and Tall from 1998, but in agreement with some other unnamed studies according to East (2007, p. 331). Therefore, East's (2007) conclusion is that there is need for further investigation in this field (p. 349).

As mentioned earlier, feedback is one of the fundamental elements in formative assessment. In her M. Paed. thesis, Milena Remis (2004) studied what kind of feedback second language students prefer in their English writing (p. 8). She collected data within second language institutions in New York (p. 31). Her study shows that students feel there is need for both feedback on grammar, on the one hand, and content and structure, on the other, and that there should be some balance between the two (p. 82). Furthermore, the study shows that teachers and students do not view error identification the same way. While teachers seem to think it works well and are inclined to use it, students seem to feel it is not a helpful

technique (p. 7), and they prefer error correction which they in fact want as much of as possible (p. 82).

2.9.2 Icelandic Research

Despite its importance, assessment has not been the target of many formal studies here in Iceland, as mentioned in chapter 1, at least not concerning methodology used in foreign language assessment. However, a few studies have been conducted in recent years on assessment methodology in educational assessment in general, as well as in English as a foreign language at the primary level. What they all found is that the most common way to assess students' progress in Icelandic schools is through paper-and-pencil tests. However, when the students' views on assessment methods were investigated, the results show that they wish for more performance-based assessment.

Two recent studies exist that investigated assessment in English in particular. Lilja Jóhannsdóttir (2008) found that the methods used for assessing English learning in the 8th grade are not in line with the teaching approaches advocated in the Curriculum Guidelines for the primary level (p. 3) even though approximately 90% of the participating teachers say that the Curriculum Guidelines influence their planning (p. 36) and almost 74% say they plan the assessment methods at the same time they plan the teaching (p. 38). According to this study, the majority of teachers use solely, or mostly, written tests for assessing their students. Other commonly used factors are listening exercises and writing tasks, whereas oral skills do not weigh as much as the others and a minority of the teachers use portfolio assessment and/or peer- and self-assessment (p. 41 – 42).

In her study, Guðfinna Gunnarsdóttir (2005) also examined, among other things, assessment methods. Her study indicates that the main emphasis in teacher assessment in the 10th grade is written tests and there is not much attention given to oral skills when it comes to

assessment (p. 46). Additionally, the study shows that it is not common for teachers to make use of peer- or self-assessment (p. 60). However, in opposition to many theorists named by Guðfinna Gunnarsdóttir (2005), the majority of teachers participating in her study grade student behaviour and effort as well as achievement (pp. 41 – 42). Moreover, according to Guðfinna Gunnarsdóttir (2005), some teachers seem to view the Curriculum Guidelines as something they do not need to consider in their choices regarding assessment (p. 64), which, along with Lilja Jóhannsdóttir's conclusions, has to be of some concern to educational authorities.

As for educational assessment methods in general, Rósa Maggý Grétarsdóttir (2007) studied secondary level teachers' views on assessment in Icelandic, where she found that teachers do not give the purposes of assessment much thought, and while they think the best way to assess students' progress is through essays or presentations of some kind, paper-and-pencil tests were the first assessment method they mentioned when asked about how they assess their students (p. 35 – 36). Erna Ingibjörg Pálsdóttir (2006) investigated methods used at the primary level. She found that the main emphasis is on paper-and-pencil tests for knowledge, in addition to some performance-based assessment in vocational subjects. Furthermore, that the standardized tests clearly affect the choice of assessment methods (p. iii). Additionally, Rúnar Sigbórsson (2008) conducted research in connection with his doctoral dissertation, where he studied the effects that standardized tests have on teaching approaches and students' learning processes in Icelandic and Science. His conclusion is that the tests have negative effects on teaching approaches, in addition to the anxiety they create for some students. Against some of the teachers' beliefs, the teaching is aimed towards the tests, and the teacher assessment emphasizes on summative assessment through written tests, often old standardized tests. There are indications of more emphasis being placed on grades than on

actual learning, and elements such as problem solving and reflection are set aside (pp. iii – iv) which is in accordance with research findings from abroad (Harlen, 2005, pp. 208 – 209).

One study was found that investigates students' views on assessment. Ragnheiður Hermannsdóttir (2008) interviewed students from two Icelandic primary schools with different assessment policies. Her conclusion is that students want formative assessment where the focus is on their work and the learning process. Also, that they feel that big exams are demotivating and do not show learners' real knowledge and abilities in the same way as projects such as essays or presentations of some kind do (p. 3).

Concerning feedback in particular, a study was conducted recently on written feedback in foreign English writing instruction. Steinlaug Sigríður Bjarnadóttir (2008) studied written feedback in English writing in the 8th and 10th grades at the Icelandic primary level where she focused on the amount of feedback given and what features of writing are addressed in the feedback (p. 6). Her findings suggest that error correction on the form is more common than consideration for the content or structure in both grades (p. 51). Also, that there was no "significant development of written feedback" (p. 48) between grades despite the fact that the students' language competence should have advanced somewhat during the two years, so more focus on content and structure might be appropriate in the 10th grade. This indicates that teachers in Icelandic primary schools are not meeting the needs of the students regarding feedback as expressed in Remis's study mentioned above.

3 The Study

This chapter discusses the methodology used in the study; the subjects and how they were chosen, what kind of data was used and how it was collected.

3.1 The Subjects

The subjects of this study were English teachers in Icelandic primary and secondary schools of varied sizes and location. According to lists on the Ministry of Education's website, there are 40 secondary schools and 176 primary schools in Iceland (Menntamálaráðuneytið, n.d.a). The aim of this study was to include a varied sample of schools from all over the country, while still having not too many given the limited time available to gather data and analyse the results. Thus, I had to resort to sampling, and made the decision to try to have equally many schools from both levels.

In choosing which schools to contact, I took the lists from the Ministry of Education's website, where schools are categorized according to their geographic location, and randomly picked 20 schools from each level so that there would be representatives from all parts of the country. Then homepages of the primary schools were studied to find out how big each of them was and after that 12 schools of varied sizes were chosen out of the 20, still making sure that they were evenly distributed throughout the country. The number 12 was chosen, in order to have 2 to 4 from each of the areas: the east, the west, the north, and the south, as well as the Reykjavík area. Finally, homepages of the secondary schools were studied to find out whether they were schools that emphasized academic study or whether they had more comprehensive programs. This was done to ensure that both types were represented in the sample, and again the same geographic criteria were used and 12 schools were picked out of the 20. However, since teachers of the primary schools were rather reluctant to participate, 8 additional primary schools were contacted later, and out of the 20, altogether 25 teachers from 9 schools agreed

to participate, whereas 36 teachers from 11 out of 12 secondary schools were willing participants. Some of the teachers answered for more than one grade level or course unit, while most of them answered only for one each.

According to the National Curriculum Guidelines for the Primary Level, English is taught from the 4th grade up, but schools also have the choice to start earlier (*Aðalnámskrá grunnskóla*, 2007, pp. 17, 47). Examining assessment at all age levels would have meant too voluminous materials, thus a choice was made to include every second year. To make sure that all the participating schools have English as a subject in the grade levels studied a choice was made to take 4th grade as the first one; hence the grade levels included in the survey were 4th, 6th, 8th, and 10th grades. Also, four course units at the secondary level were chosen out of various possibilities. Since the sample schools vary in size they also have varied course units to offer, thus their homepages were consulted in order to choose course units they all teach. As the 10th grade was included in the primary level section, the first course unit at the secondary level was left out, but since it varies between schools whether ENS 202 or ENS 203 is the second course unit offered, I decided to ask for information on either ENS 202 or ENS 203 and then view them as one in analysing the results. Thus the course units chosen at the secondary level were: ENS 202/203, ENS 303, ENS 403, and ENS 503.

3.2 The Questionnaires

As a result of the decision to include schools from all over the country in the survey, data was collected through questionnaires. In deciding what kind of questions to include, I looked back to my own experience in answering questionnaires in my work. Teachers are often asked to take time from their busy schedules to fill out surveys of various kinds, so questionnaires that take a long time to answer are rather unpopular among many of them. Knowing this, I decided to use mainly closed items in the questionnaires, though with the

possibilities for participants to add information where they needed, even though open ended questions are more likely to give the exact answers that participants want to provide (Nunan, 1992, p.143). This was believed to make the questionnaires easy to answer and encourage participation. Also, questionnaires from previous international and Icelandic studies, discussed in chapter 2.9, were viewed and the questions used here reflect those as well as the theory discussed in chapter 2. In retrospect, the questionnaires have some faults that a more experienced researcher would have been able to avoid. Examples of this will be pointed out in chapter 5.

The questionnaires were sent to the participating teachers via e-mail, using the University's K2 survey, and the answers were not traceable to participants. The first questions were about the school and the grade level/course unit in question. Then there were two questions regarding the teachers' own education and English teaching experience. The longest section of the questionnaires was about methods used for assigning final grades, and there teachers were asked whether they use tests, written and/or oral; performance-based assessment, written and/or oral work; portfolio assessment; self- assessment; and peer-assessment. Furthermore there were questions on what types of written tests are used, whether there are time limits for written tests, what kinds of tasks are used in the various types of assessment, what language skills are assessed in each way, and total percentage of each method in the final grade. Finally there were four questions on whether and then what type of formative assessment is practiced, and two questions on how final assessment results are presented to students and, when appropriate, their parents. The questionnaires can be viewed in the Appendix.

4 The Results

Data analysis is based on statistical information from the questionnaires. Out of the 36 teachers from 11 secondary schools who agreed to participate, 24 answered the questionnaire. This is 67% participation which is acceptable. As for the primary level, altogether 25 teachers from 9 schools agreed to participate and 20 of them answered, resulting in 80% participation. However, all in all, the participating teachers from the secondary level are a larger percentage of the whole secondary level than the participating primary level teachers are for the primary level and this needs to be kept in mind when viewing the results.

All in all, answers came from 11 teachers concerning ENS 202/203, 10 teachers regarding ENS 303, 8 teachers regarding ENS 503, but only 5 concerning ENS 403. From the 9 participating primary schools, answers came regarding eight 10th grades, seven 8th grades, five 6th grades, but only three 4th grades.

4.1 Teachers' Education and Experience

One of the aims of the study was to find out whether there is any noticeable connection between the choice of assessment methods and teachers' own education level and/or English teaching experience. In order to find out, the second section of the questionnaires included questions about teachers' education and experience which also gave information about the distribution of age and school levels between participants. In this chapter, the outcome of the questionnaires regarding how teachers' education level and experience is divided between age and school levels will be shown, while the second aspect, whether there is any noticeable difference in assessment method choice, will be discussed in chapter 4.2.1.

Table 1. Secondary level teachers' education and English teaching experience

Secondary level course units <i>N</i> = 34	Teachers' education				English teaching experience			
	English degree/certified teacher <i>N</i> = 19	English degree/not certified teacher <i>N</i> = 2	No English degree/certified teacher <i>N</i> = 0	Other <i>N</i> = 3	< 5 years <i>N</i> = 1	5 – 10 years <i>N</i> = 5	11 – 20 years <i>N</i> = 4	> 20 years <i>N</i> = 14
ENS 202/203	8	2		1	1	6	1	3
ENS 303	9	1			1	1	2	6
ENS 403	4	1				2	1	2
ENS 503	4	2		2	1	2	1	4

Table 2. Primary level teachers' education and English teaching experience

Primary level classes <i>N</i> = 23	Teachers' education				English teaching experience			
	English degree/certified teacher <i>N</i> = 5	English degree/not certified teacher <i>N</i> = 0	No English degree/certified teacher <i>N</i> = 11	Other <i>N</i> = 4	< 5 years <i>N</i> = 11	5 – 10 years <i>N</i> = 6	11 – 20 years <i>N</i> = 2	> 20 years <i>N</i> = 1
4 th grade			3		2	1		
6 th grade	1		3	1	2	2	1	
8 th grade	1		4	2	3	3	1	
10 th grade	3		3	2	4	2	1	1

As can be seen in table 1, a majority of the secondary school teachers who answered have a degree in English as well as a teaching certificate, altogether 19 or 79%, whereas only 2 have a degree in English without a teaching certificate and 3 have no degree in English. Also, a majority of them have extensive, or quite extensive, English teaching experience; 58% over 20 years and 17% 11 – 20 years. The three who do not have a degree in English have all got over 20 years of experience, along with one of the two who do not have a teaching

certificate. However, when it comes to the primary schools (see table 2), only 5 out of 20, or 25%, have a teaching certificate along with a degree in English, either B.Ed. or B.A., whereas 11 have teaching certificates without any university education in English and 4 do not have a university degree in English nor a teaching certificate. Additionally, a slight majority of the primary school teachers only have an experience of less than 5 years of teaching English, altogether 11 or 55%, while only one has an experience of over 20 years and two of 11 – 20 years. Furthermore, the teachers who teach the young learners in the 4th and 6th grades all have teaching certificates but no university education in English, except for one.

4.2 Summative Assessment

The biggest part of the questionnaires was about assessment methods regarding summative assessment; altogether 36 questions in section 3 (see Appendix).

4.2.1 Assessment Methods

Table 3. Assessment methods used for summative assessment

Assessment methods	4 th grade <i>N</i> = 3	6 th grade <i>N</i> = 5	8 th grade <i>N</i> = 7	10 th grade <i>N</i> = 8	ENS 202/203 <i>N</i> = 11	ENS 303 <i>N</i> = 10	ENS 403 <i>N</i> = 5	ENS 503 <i>N</i> = 8
Written tests	2	5	7	8	11	9	5	5
Written assignments		3	7	8	11	10	5	8
Oral tests	1	3	4	6	7	7	4	7
Oral presentations			5	7	7	6	3	8
Self-assessment	1	2	3	3	5	4	1	4
Peer-assessment			3	4	4	4	1	4
Portfolio	2	2	2	1	9	7	2	4
Other	1		1		4	1	2	4

Written tests are the main assessment method used by the teachers in this study, both at the primary level and in all of the course units except ENS 503 (see table 3). Even two-thirds of the young learners in the 4th grade at the primary level are assessed with written tests despite the Curriculum Guidelines' disapproval of formal assessment in the 4th grade (*Aðalnámskrá grunnskóla*, 2007, p. 17). However, generalization here is impossible since the participating 4th grade teachers were so few.

According to the teachers' answers, written tests, when used, count for 20 – 100% of the final grade at the secondary level and 25 – 75% at the primary level. However, in all cases other assessment methods are also used and the results of those count for some proportion of the final grade so at least one of the secondary level teachers has misunderstood the question. When examined closer, 30 of the 34 discussed course units at the secondary level have written tests as one of the methods to decide students' final grade. In 25 of the cases they count for 40 – 60% of the final grade. At the primary level, written tests are used in 22 of the 23 classes and in 13 cases they count for 40 – 60% of the final grade.

Where written tests are used, there are time limits in all cases except one at the secondary level and all cases except four at the primary level. When asked to explain why and how the limits are decided, teachers gave various answers. At the secondary level time limits are often decided by the school or department, but another commonly named reason was restrictions of timetable when tests are taken in lessons. At the primary level, timetable restrictions were the most common answer. In explaining how the length of tests is decided in relation to the allotted time, some mentioned experience while one said that he/she takes the tests to figure out time needed. Three teachers at the secondary level and two at the primary level mentioned additional time for students with special needs, like dyslexic students, and one at each level said that the time limits were flexible. A few of the teachers explicitly mentioned that there is ample time for all within the time limits, but one at each level said that

when tests are taken in lessons, there is the possibility of an additional 15 minutes in the break for those who need it, implying that the allotted time is not always enough for everybody.

Table 4. Types of written tests used

Types of written tests	4 th grade <i>n</i> = 2	6 th grade <i>n</i> = 5	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 9	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 5
Traditional tests	2	5	7	8	11	9	3	3
Allowing some means of aids e.g. dictionaries	2	1	3	4	3	6	3	4
Cooperation tests			2	2	3	1	1	1
Tests taken in more than one day, allowing students to take questions home			1	1	2	1	1	2
Take-home tests	1		3	5	1	1		1

Those teachers who use written tests were asked what types of tests they use and, as can be seen in table 4, traditional tests are most frequently practiced at both levels. However, other types are also used and tests allowing some means of learning aids are the most common of those at both levels. Noticable is also that take-home tests are quite common at the upper primary level. When the answers are studied closer, it can be seen that in 59% of the primary level classes, where written tests are applied, and 60% of the secondary level course units, teachers use more than one type of written tests, meaning that there is made use of at least two tests. Also, in 36% of the same primary level classes and 23% of the same secondary level course units three or more types of written tests are used.

Even though written tests are the main assessment method, according to this study, other methods are also frequently used. Written assignments are, for example, used by all the teachers at the secondary level and the upper primary level, as well as some of the teachers

who teach 6th grade (see table 3). At the secondary level, these assignments count for 15 – 40% of the final grade if written tests are used as well, but 50 – 80% when there are no written tests. At the primary level, these assignments count for 10 – 40% of the final grade. When the answers are examined further, it can be seen that in 89% of the primary level classes, where written assignments are used, and 97% of the secondary level course units, more than one type of written assignments are used, and in over 70% cases at both levels three or more different types of assignments are used.

At the secondary level, the majority of the teachers use formal assessment of oral skills in the form of oral tests and/or oral presentations (see table 3). However, in two of the ENS 202 and two of the ENS 303 course units there is no formal assessment of oral skills despite the Curriculum Guidelines' emphasis on the practice and assessment of all language skills according to goals in those course units (*Aðalnámskrá framhaldsskóla*, 1999, pp. 32 – 34; *Aðalnámskrá framhaldsskóla*, 2000, pp. 11 – 13). The oral tests count for 5 – 35% of the final grade and the oral presentations 2.5 – 30%. It is interesting that the highest percentage is in ENS 503 where the Curriculum Guidelines do not mention oral skills particularly since that course unit emphasizes on writing skills (*Aðalnámskrá framhaldsskóla*, 1999, p. 38). A closer study of the answers reveals that in 80% of the secondary level course units, where oral tests are used, teachers use more than one type of task in oral tests, which can mean that there are more tests than one. Also, in 48% of the same secondary level course units three or more different types of tasks are used in oral tests. Furthermore, in 71% of the secondary level course units, where oral presentations are applied, more types than one are used, and in 38% of the same secondary level course units three or more different types of oral presentations are used.

At the primary level, no formal assessment of oral skills is performed in 7 out of 23 classes (see table 3). Two of the 8th grade teachers and one of the 10th grade teachers do not

use any form of assessment of oral skills, despite the Curriculum Guidelines' emphasis on taking all language skills into account when assessing students' abilities (*Aðalnámskrá grunnskóla*, 2007, p. 18). However, despite the Curriculum Guidelines' disapproval of formal assessment in 4th grade (*Aðalnámskrá grunnskóla*, 2007, p. 17), one of the 4th grade teachers does apply oral tests, while one of them assesses oral skills through in-class discussions, and one of them has no assessment of oral skills. Only three of the five 6th grade teachers use oral tests and none of them uses oral presentations for assessing oral skills even though those skills are heavily emphasized in the Curriculum Guidelines for this level (*Aðalnámskrá grunnskóla*, 2007, p. 18). The oral tests, when used at the primary level, count for 5 – 30% of the final grade, the highest percentage being in the 6th grades, and the oral presentations count for 5 – 20% with the highest percentage here being in the 10th grades. When examined closer, it can be seen that in 86% of the primary level classes, where oral tests are used, teachers use more than one type of task in oral tests, which indicates that there are more tests than one. Also, in 57% of the same primary level classes three or more different types of tasks are used in oral tests. Furthermore, in 92% of the primary level classes, where oral presentations are applied, more types than one are used, and in 58% of the same primary level classes three or more different types of oral presentations are used.

Self-assessment, according to table 3, is used in less than half of the cases in this study. At the secondary level, 14 out of 34 course units have self-assessment as 2.5 – 10% of the final grade, whereas 9 out of 23 classes at the primary level do so. This is not in line with the emphasis of the Curriculum Guidelines for the primary level which actually encourage the use of self-assessment (*Aðalnámskrá grunnskóla*, 2006, p. 17). Peer-assessment is even less practiced than self-assessment at the primary level; only half of the 10th grade teachers, 43% of the 8th grade teachers and none of the younger grades' teachers make use of peer-

assessment. On the other hand, peer-assessment and self-assessment are almost equally practiced at the secondary level.

Another assessment method advocated in the Curriculum Guidelines for the primary level is portfolio assessment (*Aðalnámskrá grunnskóla*, 2007, p. 13). When the results in this study are viewed however, English teachers at the primary level do not seem to have adopted that method (see table 3). Only in 7 out of 23 classes, portfolio assessment of some kind is used, counting for 5 – 30% of the final grade. It seems to be more common amongst those who teach the younger students, since it is only applied in one of eight 10th grades and two of seven 8th grades. This is particularly interesting when the results from the secondary teachers are viewed. There, portfolio assessment is applied in 22 out of 34 course units even though this method is not particularly advocated in the Curriculum Guidelines for the secondary level.

Some of the teachers named other aspects or assessment methods used for assigning the final grade. The teacher of a 4th grade who does not use written tests, nor any of the other methods discussed in this study apart from portfolio assessment, talks about assessing students' talk in class, where they are encouraged to use previously practiced vocabulary. The other primary level teacher who mentions other aspects than discussed in the study, talks about assessing students' in-class effort and homework. That is also referred to by three secondary level teachers. Finally, one of the ENS 503 teachers uses an online portfolio which counts for 100% of the final grade.

Table 5. Connection between assessment methods and teachers' education and experience

Assessment methods	Teachers' education				English teaching experience			
	English degree/certified teacher <i>n</i> = 24	English degree/not certified teacher <i>n</i> = 2	No English degree/certified teacher <i>n</i> = 11	Other <i>n</i> = 7	< 5 years <i>n</i> = 12	5 – 10 years <i>n</i> = 11	11 – 20 years <i>n</i> = 6	> 20 years <i>n</i> = 15
Written tests	21	2	10	5	11	10	6	12
Written assignments	22	2	6	7	9	9	6	15
Oral tests	16	2	5	4	7	7	4	11
Oral presentations	12	2	4	4	7	7	1	9
Self-assessment	7	1	5	1	4	7	2	1
Peer-assessment	7	1	3	0	2	5	2	2
Portfolio	13	1	3	1	6	4	5	6
Other	3	2	3	0	2	3	0	2

When the connection between chosen assessment methods and teachers' education level or English teaching experience is examined, there are no obvious correlations (see table 5). However, there are some indications worth noticing. With regard to education, those who have a B. Ed. degree without specialization in English are less likely to apply written assignments, oral tests, and oral presentations than the others. This, however, can also be explained with the fact that 55% of those teachers teach the youngest students, the 4th and 6th graders. Additionally, the certified teachers who have specialized in English are more likely to apply portfolio assessment than the others. When English teaching experience is viewed, those with the least experience are the least likely to apply oral tests, while those with the most experience are the most likely. Also, those with the most experience are the least likely to apply self-assessment and peer-assessment, and those with 11 – 20 years worth of experience are the most likely to apply portfolio assessment but the least likely to apply oral

presentations. However, due to how small the sample is in this study, it can not be said with certainty that any of these differences are statistically significant, but the results show some indications.

4.2.2 Types of Tasks Used

In order to view how varied the assessment methods actually are and whether teachers follow suggestions given in the Curriculum Guidelines, the questionnaires included some questions regarding different types of tasks in written and oral tests, different types of assignments used, both written and oral, and different types of form of portfolio assessment. The answers to these questions can be seen in tables 6 – 10.

Table 6. Task types in written tests

Task types	4 th grade <i>n</i> = 2	6 th grade <i>n</i> = 5	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 9	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 5
Multiple-choice items	2	1	6	7	10	7	4	4
Open questions		3	6	7	11	8	5	5
Gap-filling items	2	3	5	7	10	8	4	3
True/false items	2	2	5	6	5	3	2	1
Matching items	2	4	5	7	9	8	1	3
Translation into Icelandic	2	3	5	3	3	3	2	
Pre-read texts	2	5	4	3	5	7	2	2
Unseen texts	1	4	7	8	11	8	4	3
Free writing		3	6	7	8	7	2	3
Translation into English			4	2	2	3	1	1
Other						1		1

The teachers were asked what types of tasks they use for the various methods. As can be seen in table 6, there is quite some variety in the tasks used in written tests. The two 4th grade teachers, who use written tests, seem to use the same kinds of tasks except one of them uses unseen texts in addition to the pre-read texts they both use. Multiple-choice items are commonly used except in the 6th grade, and almost everyone uses open questions as well. True/false items are more popular at the primary level than the secondary level, but mainly it is not as commonly used as many of the other items. The same goes for translation into Icelandic, and translation into English is the type of task that is used the least. Two secondary level teachers mentioned other things than those in the questionnaire; literary interpretation and data gathering from the Internet to use in assignments.

Table 7. Types of written assignments

Task types	4 th grade <i>n</i> = 0	6 th grade <i>n</i> = 3	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 10	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 8
Retelling		1	4	4	5	3	1	2
Summaries			4	6	9	5	4	6
Book reports			4	6	9	8	4	5
Essays			6	7	10	9	3	7
Free writing		2	6	8	8	6	2	3
Translation into English		1	5	4	2	1	1	1
Other			1	2				1

Table 7 shows the use of varied written task types. Free writing and essays are the task types that are the most common at the primary level, while essays, book reports and summaries are the most frequently used at the secondary level, which is in accordance with the suggestions made in the Curriculum Guidelines (*Aðalnámskrá framhaldsskóla*, 1999, p.

21). Translation into English is quite commonly used at the upper primary level, but not so much at the secondary level. This type of practice is discussed in the Curriculum Guidelines for the primary level, and there it is explicitly described as not being a task that practices writing skills, only grammar and vocabulary (*Aðalnámskrá grunnskóla*, 2007, p. 10). One teacher at the secondary level mentioned other tasks such as plays, poems, interpretation and considering of various texts. Also, three teachers at the primary level referred to other tasks than named in the questionnaires; the making of a dictionary, scripts for short films, lyrics for music they compose, interviews, and stories from series of pictures.

Table 8. Task types in oral tests

Task types	4 th grade <i>n</i> = 1	6 th grade <i>n</i> = 3	8 th grade <i>n</i> = 4	10 th grade <i>n</i> = 6	ENS 202/203 <i>n</i> = 7	ENS 303 <i>n</i> = 7	ENS 403 <i>n</i> = 4	ENS 503 <i>n</i> = 7
Pre-read text discussed	1	2	3	5	7	6	3	7
Film or TV-show discussed			3	6	4	2	2	4
Pictures viewed and discussed	1	1	3	3	2	1	2	1
Discussion on daily life	1	2	3	4	3	3	1	5
Discussion between students				2	2	2	1	3
Other			2	2	1	2	1	2

As can be seen in table 8, various types of tasks are used in oral tests. Discussing a pre-read text is by far the most frequently used task at the secondary level, used by almost all the teachers who have oral tests. Discussing daily life and films or TV-shows is also often used, and these three task types are the most commonly used at the primary level as well. Discussing pictures is more commonly used at the primary level, while discussion between students is more frequent at the secondary level. One teacher at the secondary level named pre-decided subjects of discussion that are in accordance with the textbook used in class, and

one referred to the acting out of a short play. Furthermore, a few mentioned that the pre-read texts used are whole novels which are discussed in the tests. The other task mentioned by primary level teachers is the description of an object students bring from home.

Table 9. Types of oral presentations

Task types	4 th grade <i>n</i> = 0	6 th grade <i>n</i> = 0	8 th grade <i>n</i> = 5	10 th grade <i>n</i> = 7	ENS 202/203 <i>n</i> = 7	ENS 303 <i>n</i> = 6	ENS 403 <i>n</i> = 3	ENS 503 <i>n</i> = 8
Presentation of read material			3	5	3	3	2	5
Presentation of film or TV-show			2	6	1	1	1	1
Presentation of own choice			4	5	3	3	1	4
Power point presentation			4	6	2	4	2	7
Other				1	3	2	1	2

As can be seen in table 9, no oral presentations of any kind are used for assessing the younger students of 4th and 6th grades. Using power point presentations seems to be a popular way at both educational levels, especially the upper primary level. The presentation of films and TV-shows is not common at the secondary level, but more so at the primary level. Other tasks named here, by the secondary level teachers, are presentations of something related to English or global culture and literature, for example famous buildings, places or people, or the African-Americans' fight for equal rights in the USA. One 10th grade teacher mentioned a fairly different type of assignment: the presentation of a self-made short film or music.

Table 10. Portfolio assessment

Portfolio assessment	4 th grade <i>n</i> = 2	6 th grade <i>n</i> = 2	8 th grade <i>n</i> = 2	10 th grade <i>n</i> = 1	ENS 202/203 <i>n</i> = 9	ENS 303 <i>n</i> = 7	ENS 403 <i>n</i> = 2	ENS 503 <i>n</i> = 4
Logbook	1	2				1		1
Process portfolio	1		2	1	9	6	2	3
Showcase portfolio						2		1
Other								1

As discussed in chapter 4.2.1, portfolio assessment is not commonly practiced at the primary level, but when used, it includes either the use of logbooks or process portfolios, or both (see table 10). At the secondary level, on the other hand, the use of portfolio assessment is more common, however, mainly the process portfolio. Only in the case of one teacher is there made use of all three: logbooks, process portfolios, and showcase portfolios. Finally, one teacher named an electronic portfolio, which seems to be of the process type.

4.2.3 Language Skills Assessed

In order to view whether all areas of the language are assessed in accordance with the emphasis of the Curriculum Guidelines, teachers were asked which skills they assess with which assessment method used, and the questionnaires named listening, speaking, reading, writing, vocabulary, grammar, and spelling. Tables 11 – 17 show what methods are used for assessing each of these skills in each of the primary level classes and secondary level course units.

Table 11. Methods used for assessing listening skills

Assessment methods	4 th grade <i>n</i> = 2	6 th grade <i>n</i> = 5	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 7	ENS 303 <i>n</i> = 7	ENS 403 <i>n</i> = 3	ENS 503 <i>n</i> = 4
Written tests	2	4	7	8	4	3	1	2
Written assignments			2	2	2	2	1	2
Oral tests	1	1	2	1	2	4	1	3
Oral presentations			1		1			1
Self-assessment		1	1			1		1
Peer-assessment			1		2	1	1	2
Portfolio		1	1	1	2	2	1	2

When assessing listening skills, primary level teachers mainly use written tests, however, some teachers also use other methods such as written assignments and oral tests (see table 11). Additionally, portfolio assessment, self-assessment, and peer-assessment are used by four teachers. At the secondary level, there is more variance in methods. However, when examined closer, teachers of only 21 out of 34 course units at the secondary level include the assessment of listening skills, while all except one at the primary level do so. This goes against the learning goals of the Curriculum Guidelines for the secondary level, where the practice and assessment of listening is emphasized for all the course units.

Table 12. Methods used for assessing oral skills

Assessment methods	4 th grade <i>n</i> = 2	6 th grade <i>n</i> = 3	8 th grade <i>n</i> = 5	10 th grade <i>n</i> = 7	ENS 202/203 <i>n</i> = 9	ENS 303 <i>n</i> = 8	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 8
Oral tests	1	2	4	6	7	7	4	7
Oral presentations			5	7	7	6	3	8
Self-assessment		2	3	3	1	1		2
Peer-assessment			3	4	2	2		3
Portfolio		1	1	1	2	2		3
Other	1							

As can be seen in table 12, oral tests and oral presentations are almost equally practiced for assessing oral skills, except in the 4th and 6th grades where oral presentations are not used and only three out of eight teachers use oral tests. Self-assessment and peer-assessment are more common here than when assessing listening, both at the primary and secondary levels. One teacher refers to a method which is not mentioned in the questionnaires: assessing in-class discussion where the students are encouraged to use previously practiced vocabulary. As with the listening skills, not all teachers assess oral skills even though the Curriculum Guidelines emphasize the practice and assessment of those in all grade levels at the primary level and all course units except ENS 503 at the secondary level. However, the roles are reversed. At the primary level the teachers of 17 out of 23 classes assess oral skills, while the teachers of 30 out of 34 course units at the secondary level do so.

Table 13. Methods used for assessing reading skills

Assessment methods	4 th grade <i>n</i> = 2	6 th grade <i>n</i> = 5	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 10	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 8
Written tests	2	5	7	8	11	9	5	5
Written assignments		1	4	3	8	7	3	7
Oral tests	1	1	2	4	6	5	3	5
Oral presentations			4	4	5	2	2	4
Self-assessment		2	2	1	3	2	1	2
Peer-assessment			3	2	3	2	1	2
Portfolio		1	2	1	8	5	2	4

Table 13 shows what methods are used for assessing reading skills. All the secondary level teachers assess reading skills in all the course units they teach, and all except one of the primary level teachers. Written tests are the most frequent method at both levels, used by almost all the teachers. At the secondary level, written assignments, oral tests, and portfolio assessment are also common, whereas these and other methods are only used by approximately one-third or less of the primary level teachers. At the secondary level, self-assessment and peer-assessment are a little more common here than when assessing speaking, but less so at the primary level.

Table 14. Methods used for assessing writing skills

Assessment methods	4 th grade <i>n</i> = 2	6 th grade <i>n</i> = 5	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 10	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 8
Written tests	2	5	7	8	11	8	5	3
Written assignments		1	7	8	11	10	4	7
Oral presentations			4	5				
Self-assessment		1	3	3	5	4	1	4
Peer-assessment			3	2	2	2		3
Portfolio		2	2	1	9	6	2	4

As for the assessment of writing, the results can be viewed in table 14. Like reading, all the secondary level teachers assess writing in all the course units they teach, and all except one of the primary level teachers. The most frequent assessment methods at the secondary level are written assignments and written tests, followed by portfolio assessment. Written tests and written assignments are also the most common at the primary level, but unlike the secondary level teachers, the primary level teachers also use oral presentations for assessing writing skills. This is interesting, but a likely explanation is that students hand in written material along with giving the presentation, thus the teacher can assess writing along with other factors. At the secondary level, self-assessment is more common for assessing writing than the other language skills, while peer-assessment is approximately as much used here as for assessing the other skills. At the primary level, however, both self-assessment and peer-assessment are approximately as infrequent while assessing writing, reading and speaking.

Table 15. Methods used for assessing vocabulary

Assessment methods	4 th grade <i>n</i> = 3	6 th grade <i>n</i> = 4	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 10	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 8
Written tests	2	4	7	8	11	9	5	5
Written assignments		2	7	8	11	10	4	8
Oral tests	1	2	3	4	6	6	4	7
Oral presentations			5	7	6	6	3	8
Self-assessment		1	3	2	4	3	1	4
Peer-assessment			3	4	1	3		3
Portfolio		2	2	1	9	5	2	4

As with reading and writing, all the secondary level teachers assess vocabulary, and all except one of the primary level teachers (see table 15). Again, written tests and written assignments are the most common ways to assess vocabulary at both levels. All teachers except one at the secondary level use written assignments and all except one at the primary level, who do assess vocabulary, use written tests. Oral tests, oral presentations, and portfolio assessment are also commonly used at the secondary level, whereas portfolio assessment is rather scarce at the primary level. At both levels, self-assessment and peer-assessment are approximately as common here as for assessing writing skills.

Table 16. Methods used for assessing grammar

Assessment methods	4 th grade <i>n</i> = 1	6 th grade <i>n</i> = 5	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 9	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 8
Written tests	1	4	7	8	11	9	5	4
Written assignments			7	7	11	8	4	7
Oral tests			2	4	6	5	3	6
Oral presentations			4	7	6	5	3	6
Self-assessment		2	2	1	3	2	1	2
Peer-assessment			3	2		2		1
Portfolio		2	2	1	9	5	1	4

Slightly fewer teachers include the assessment of grammar when calculating final grades, or teachers of 33 out of 34 course units at the secondary level and teachers of 21 out of 23 classes at the primary level (see table 16). Two of the 4th grade teachers, and one of the ENS 303 teachers do not include grammar in their summative assessment. The pattern of frequency for each method is very similar to the one for assessing vocabulary. The noticeable changes are that oral tests are not so commonly used at the primary level, and the same can be said about self-assessment and peer-assessment at the secondary level.

Table 17. Methods used for assessing spelling

Assessment methods	4 th grade <i>n</i> = 1	6 th grade <i>n</i> = 5	8 th grade <i>n</i> = 7	10 th grade <i>n</i> = 7	ENS 202/203 <i>n</i> = 11	ENS 303 <i>n</i> = 10	ENS 403 <i>n</i> = 5	ENS 503 <i>n</i> = 7
Written tests	1	5	7	7	10	8	5	3
Written assignments		2	7	7	11	10	5	7
Oral presentations			3	3	1	1	1	3
Self-assessment		1	1	1	4	3	1	3
Peer-assessment			1		1	2		2
Portfolio		2	2	1	8	5	1	4

Finally, in this section, the questionnaires viewed what methods are used for assessing spelling (see table 17). Again most of the teachers assess spelling, the same number at the secondary level as for assessing grammar, but slightly fewer at the primary level, or the teachers of 20 out of 23 classes. As with grammar, two of the 4th grade teachers do not include assessment of spelling in their summative assessment; neither does one 10th grade teacher nor one teacher of ENS 503. Once again, written tests and written assignments are the most frequent methods used. Portfolio assessment is also common at the secondary level. Self-assessment and peer-assessment are scarce at the primary level, but approximately as common at the secondary level as for writing and vocabulary. Interestingly, spelling is assessed through oral presentations in six secondary level course units and six primary level classes. A likely explanation is that following the presentation students hand in written texts or power point slides where spelling, among other factors, is considered.

4.3 Formative Assessment

One of the aims of this study was to find out whether formative assessment is practiced in Icelandic primary and secondary schools. Table 18 shows the frequencies of

different forms of formative assessment, according to the participating teachers. The questions answered are in section 4 of the questionnaires (see Appendix).

Table 18. Formative assessment

Formative assessment	4 th grade <i>n</i> = 1	6 th grade <i>n</i> = 4	8 th grade <i>n</i> = 5	10 th grade <i>n</i> = 8	ENS 202/203 <i>n</i> = 8	ENS 303 <i>n</i> = 7	ENS 403 <i>n</i> = 2	ENS 503 <i>n</i> = 5
Written feedback on assignments		3	5	8	8	7	2	5
Oral feedback on assignments	1	2	4	7	7	5	2	3
Oral feedback in discussion	1	3	4	6	4	2	1	2
Teacher-student interviews	1	1	3	7	6	4	2	4
Interviews in small groups	1		2	2	1	1		
Student participation in goal-setting		1	1	1	2	1		1
Formative use of summative tests			2	2	2	2	1	1
Other				1				

According to this study, formative assessment is more frequently used at the primary level than the secondary level. Altogether, 78% of the primary level classes have formative assessment as part of their learning environment, while only 65% of the secondary level course units do so. At the secondary level, all of the teachers who use formative assessment say they use written feedback on assignments, and a good majority also uses oral feedback on assignments and teacher-student interviews. Also, oral feedback in discussion is fairly common. However, interviews in small groups are rarely used and formative use of summative tests along with student participation in goal-setting are only used in 27% and 18% cases respectively. At the primary level, written feedback on assignments is also the most common form, closely followed by oral feedback on assignments and oral feedback in discussion. Two-thirds practice teacher-student interviews while less than one third uses

interviews in small groups. The least used forms at the primary level are formative use of summative tests and student participation in goal-setting.

One of the 10th grade teachers mentioned that when students hand in assignments, they get feedback and possibilities to correct mistakes, and then hand in again. One teacher at the primary level explained the formative use of summative tests to be in the way that when some aspects do not come out well on tests, they are given more attention in class and practiced more thoroughly in hope of student improvement, while another teacher said that all tests are gone through with the students after the marking. One teacher at the secondary level stated that if students do not succeed as well in tests as expected, they get a teacher-student interview where a plan is made for further studies. Also, in some cases students are given the chance to improve their grade. Two other teachers said students are given the opportunities to improve their work after having seen the teacher's marking, thus raise their grades.

4.4 Presentation of Assessment Results

The final questions of the survey asked about how final assessment results are presented to students and, where appropriate, parents.

Table 19. Presentation of assessment results

Presentation of assessment results	4 th grade <i>N</i> = 3	6 th grade <i>N</i> = 5	8 th grade <i>N</i> = 7	10 th grade <i>N</i> = 8	ENS 202/203 <i>N</i> = 11	ENS 303 <i>N</i> = 10	ENS 403 <i>N</i> = 5	ENS 503 <i>N</i> = 8
Grades (number/letter/word)	1	1	1	2	6	8	3	3
Written comments								
Grades + written comments	2	2	2	4	2	1		1
Grades + oral comments		1	1		1			
Grades + written and oral comments		1	3	2	2	1	2	4

As can be seen in table 19, final assessment is in all cases presented in grades; either number, letter or word grades. However, a slight majority, 56%, also gives comments; either written, oral or both. At the secondary level, 59% only present grades and no comments along with them, whereas at the primary level, that applies to only 22%. Merely oral comments with the grades are not very common, neither at the primary nor secondary level, whereas grades with written comments are the most common form at the primary level. Even though the Curriculum Guidelines for the primary level advocate oral and written comments for the youngest students, in the 4th and 6th grades (*Aðalnámskrá grunnskóla*, 2006, p. 18; *Aðalnámskrá grunnskóla*, 2007, pp. 17 – 18), this is not so in all the cases in this study. One fourth of the teachers who teach these classes only present a grade, no comments.

5 Discussion

The main aim of this study was to describe assessment practices in Icelandic primary and secondary schools. The description is based on teachers' reports, not an observation of their practices; hence, there is the possibility of misinterpretations of some kind, or even over- or under-reporting by the teachers. This study shows variety in assessment methods even though there is a tendency for overemphasis on written tests. The results of this study are thus in harmony with the previous Icelandic studies of assessment methods practiced at the primary level, in that written tests are the predominant assessment method used. However, even though too many teachers exclude oral and listening skills from their assessment practices, many teachers use performance-based, written and oral assessment of various kinds. Also, self-assessment, peer-assessment, and portfolio assessment are practiced even though in considerably fewer cases than tests and performance-based assignments. Finally, in accordance with the last three methods being used by only some participants, formative assessment seems to be ignored by too many teachers at both levels, especially the secondary level.

5.1 Differences Regarding Teachers' Education and Experience

When deciding to ask about teachers' education level and English teaching experience, the main reason was to try to find out whether there is any noticeable difference in choice of assessment methods regarding those elements. However, in viewing the results of the study, some other interesting factors appeared. One is how education and experience vary between school levels. While the majority of secondary level teachers have extensive, or fairly extensive, English teaching experience, the opposite is the case at the primary level. Also, there is a huge difference regarding the education. While a majority of the secondary level teachers have a degree in English as well as a teaching certificate, only one fourth of the

primary level teachers have some English education at university level. The difference in education can be explained with different requirements set for the teaching positions at each level, whereas the difference in experience is more difficult to explain.

Another noteworthy aspect is that at the primary level, the mainstream policy seems to be to have the classroom teacher teach English to the youngest students, even though they have no university education in English or the teaching of foreign languages. This is a somewhat distressing fact, as this is where the foundation of language study is laid and apart from a solid knowledge of the language itself and teaching approaches, teachers actually need to have some knowledge of second language acquisition in order to manage the teaching and assessing well (Zangl, 2000, p. 256). However, the survey had no questions about whether those respondents who do not have university level English education had taken courses on the teaching of English as a foreign language. Thus, it would be interesting to study further what sort of education the teachers who teach beginners have, with special emphasis on whether they have any knowledge of second language acquisition and pedagogy and language assessment in particular.

Regarding the choice of assessment methods, however, there do not seem to be any definite distinctions that can solely be explained with either different education or experience. Some of the indications that the results give can also be explained with the age of the teachers' students, and the difference can not be said to be statistically significant.

5.2 Written Tests

The fact that written tests seem to be the leading assessment method at both school levels, gives reason for further investigation of the tests. According to the teachers' answers in this study, traditional paper-and-pencil tests are the most commonly used, indicating that the behaviourist theories of learning, from the early twentieth century, are still controlling the

testing of English learning in Iceland despite the emphasis that theorists have placed on cognitive, constructivist, and socio-cultural theories in recent years. Still, approximately half of the teachers at both levels also administer tests where some means of learning aids are allowed, for example dictionaries. Also, 9% of the primary level classes and 20% of the secondary level course units, which are assessed with written tests, are in some cases assessed with tests taken in more than one day where the students can take questions home between days to study further. And take-home tests are sometimes used for 41% of the primary level classes and 10% of the secondary level course units. According to East (2007), all this indicates that constructivist theories are also underlying, at least where some of the teachers are concerned (pp. 332 – 333). Furthermore, socio-cultural theories seem to affect the choice of approximately 20% of the teachers at both levels, who use written tests, since those apply cooperative tests in at least some cases.

As this study shows, time limits are the standard procedure when it comes to written tests. It is important to bear in mind that the allowed time has to be adequate so the inferences of the results of the tests can be valid (Gronlund, 2003, p. 205). However, some of the answers in this study imply that there is not always allotted enough time. According to Bachman and Palmer (1996) these are called ‘speeded tests’ and are designed in a way that makes sure that not all students manage to complete all the tasks, whereas tests that are meant to give enough time for all students are called ‘power tests’ (p. 51). The answers in this study are not thorough enough to see whether tests in Icelandic schools are created as speeded or power tests.

This discussion on the practice of written tests in Icelandic schools shows that there is a need for a further study of used tests. A study that focuses on Bachman and Palmer’s philosophy of language testing (see chapter 2.4.1), would be interesting and might lead to informed knowledge of the tendency in Icelandic schools and how reliable the test results can

actually be considered. Since there are no longer external standardized tests for the Icelandic school system, the results of classroom-based assessment play an even more important role than before. They affect students' possibilities for further studies, whether at the secondary level after finishing primary school, or at the tertiary level after finishing secondary school, so a solid knowledge of what is assessed, and how, must be of great importance for the educational system on the whole.

5.3 Alternative Assessment Methods, In-class Effort and Homework

Currently, many researchers consider the so-called alternative assessment methods to be important within the field of assessment. Among others, McKay (2006) has pointed out how valuable portfolio assessment can be since it involves student self-reflection, which again leads to more autonomous students. Also, portfolio assessment can have positive effects on the activity choice in the classroom and make the connection between learning activities and assessment clearer. Additionally, it can improve teacher professionalism and involve parents in the assessment process. Finally, it has the benefit of making assessment a teaching strategy which results in assessment serving to enhance learning (pp. 159 – 160).

In accordance with this, the National Curriculum Guidelines for Foreign Languages at the Primary Level advocate the use of portfolio assessment and explicitly mention logbooks as valuable aid in the learning and assessment processes (*Aðalnámskrá grunnskóla*, 2007, pp. 13, 18). Thus, it is disappointing to see that so few teachers at that level actually practice this assessment method; only 7 teachers, of the 20 who participated, claim they use portfolio assessment, and only 3 of them use logbooks. Since heavy emphasis on the importance of student self-reflection and autonomy can be found in the Curriculum Guidelines for the primary level (*Aðalnámskrá grunnskóla*, 2007, p. 13, 18), it is hard to understand why teachers do not choose portfolio assessment. Also, it is interesting that in spite of no particular

mentioning of this method in the Curriculum Guidelines for the secondary level, teachers there seem to have adopted this method in more quantities than at the primary level. However, it is noticeable that the majority of secondary level teachers, who apply portfolio assessment, only use the process portfolio.

The questionnaires did not include specific enough questions on this subject to get answers to why this method is not chosen by more primary level teachers, nor why it seems to be more popular at the secondary level. Neither does this study provide information on how the particular forms are practiced by the teachers. According to Gronlund (2003), it is not enough to just collect samples of work into a folder; portfolio assessment involves much more than that (p. 159). There is the important student self-reflection and goal-setting, there is the urgency of systematic sample collection and marking using predetermined criteria, and there is the essential feedback from the teacher (McKay, 2006, pp. 159, 161). Thus, a further study on this would be interesting. It has been claimed that portfolio assessment is quite time consuming (Gronlund, 2003, p. 159), and that could be one of the factors that affect the teachers in their choice. Also, teachers might lack training in how to apply this assessment method.

Performance-based written assessment is commonly used at both levels and often counts for quite a lot of the final grade, while performance-based oral assessment is less common. At the primary level, 61% of the classes are assessed with oral tests, and 52% with oral presentations, whereas one teacher uses in-class discussion as basis for assessment of oral skills. At the secondary level, this type of assessment is more common than at the primary level, 74% of the course units are assessed with oral tests and 71% with oral presentations.

As for self-assessment and peer-assessment, those have not been adopted by the majority of teachers according to this study, which is in accordance with the previous studies conducted by Guðfinna Gunnarsdóttir (2005) and Lilja Jóhannsdóttir (2008). According to

this study, only approximately 40% of participating teachers at both levels apply self-assessment. This is especially noteworthy considering the emphasis placed by the Curriculum Guidelines on that type of assessment, formative assessment, and learner autonomy at the primary level, as well as on formative assessment at the secondary level. According to Black et al. (2003), self-assessment is a fundamental aspect of formative assessment (p. 18), and McKay (2006) has claimed that self-assessment is an effective way to get students to understand their studies better and become more autonomous (pp. 165 – 166). Hence, the lack of practicing self-assessment in Icelandic schools is a matter of concern for the educational system. As for peer-assessment, it is almost equally practiced at the secondary level as self-assessment, while only applied in 30% of the primary level classes. In a way this is understandable, since peer-assessment does have some limitations for young learners (McKay, 2006, p. 166). However, peer-assessment is considered to be equally effective as self-assessment and have the same or similar benefits for the students (James & Pedder, 2006, p. 28; Weeden, Winter, & Broadfoot, 2002, p. 89), which leads to the same conclusion as for self-assessment, that the lack of its practicing in Icelandic schools is a matter of concern.

The teachers were asked whether they include any other factors than students' achievement of language skills when assigning the final grade. Not many responded to that question. However, one primary level teacher and three secondary level teachers said they include in-class effort and homework. This is noticeably different from the study conducted by Guðfinna Gunnarsdóttir (2005), where 9 of 11 teachers include homework and 6 of 11 include in-class effort (p. 41). Furthermore, none of the participating teachers in this study mentioned aspects such as behaviour, attendance, motivation or initiative, which were all included by 6 to 7 teachers in Guðfinna Gunnarsdóttir's (2005) study (p. 41).

5.4 Variety in Assessment Methods

The Curriculum Guidelines for both levels advocate continuous assessment with the argument that it is in harmony with language learning (*Aðalnámskrá grunnskóla*, 2007, pp. 13, 18; *Aðalnámskrá framhaldsskóla*, 1999, pp. 15, 22). According to this study, teachers at both primary and secondary levels seem to have adopted continuous assessment. Unfortunately the questions in the questionnaires on this subject were not specific enough to give concrete answers on how many actually practice continuous assessment, but it is safe to say that there are strong indications that continuous assessment is being practiced by most, if not all, teachers.

Furthermore, the Curriculum Guidelines for both levels advocate diversity in choice of assessment methods and tasks (*Aðalnámskrá grunnskóla*, 2007, p. 13; *Aðalnámskrá framhaldsskóla*, 1999, p. 21). All the participating teachers, at both levels, use more than one method of assessment, and the majority of secondary level teachers use four or more methods while the majority of primary level teachers use three or more methods. However, once again, the questions of the questionnaires do not give specific enough answers. There were not questions on how many tests or assignments are used in each class or course unit, nor how often self-assessment or peer-assessment are practiced. Thus, the information that this study gives on the actual variety of tasks is limited. However, one can infer from the answers to questions on what types of written tests are used, what types of tasks are used in oral tests, and what types of written and oral assignments are used. According to this study, the majority of classes at both levels are assessed with more than one written test. Also, a good majority of classes at both levels are assessed with three or more different types of written assignments. A good majority of classes at both levels are assessed with various types of tasks in oral tests, whether that means there are more than one tests used or whether there are used more than one type in each test. Finally, this study shows that a good majority of classes at both levels

are assessed with more than one type of oral assignments. Hence, the results of this study show strong indications to variety in assessment methods at both primary and secondary levels in Icelandic schools.

5.5 Assessment of All Language Skills

The Curriculum Guidelines for both levels highlight the importance of assessing all language skills in accordance with the emphasis laid in the instruction, thus embrace the communicative approach to language instruction, which, according to Bachman and Cohen (1998), was developed in the 1980s as a result of current second language acquisition theories (pp. 4 – 5). Still there are teachers who do not consider listening or oral skills when assessing their students and that has to be of concern. Teachers of only 62% of the course units at the secondary level assess listening skills, and only 74% of the primary level teachers assess oral skills. These results are in accordance with research results discussed by Rea-Dickins (2007), which show that speaking and listening are often left un-assessed where classroom-based assessment is concerned (p. 513). Regarding the assessment of oral skills at the primary level, these results are also similar to the results of Lilja Jóhannsdóttir's (2008) study, where 23% of participating 8th grade teachers do not use oral tests and 13% do not use oral assignments (p. 41). However, they are far better than those found in the study of Guðfinna Gunnarsdóttir (2005), where only three of eleven 10th grade teachers assessed oral skills (p. 46). Furthermore, the difference between school levels here is noteworthy. All teachers except one at the primary level assess listening while teachers of only 62% of the secondary level course units do so, whereas teachers of 88% of the secondary level course units assess oral skills while only 74% of the primary level teachers do so.

Unfortunately, there was no question in the questionnaires on why teachers do not assess all the skills that are, or should be, a part of the learning process, so there is no

explanation to be found here. Whether it is because the teachers also ignore these skills in the teaching or whether it is only in the assessing is left unanswered. However, it is not unlikely that, at least in some cases, this is because the teachers are not familiar enough with possible ways to assess listening or oral skills. Also, there is the danger that when there is shortage of time, the ‘easiest’ way to cover many things in a short period of time will be chosen, leaving oral assessment out since its practice undeniably takes more time than written assessment.

It would be worth while to study this further. Why do not all English teachers see the importance of assessing all language skills? What can be done to change this, so that the assessment results that students get actually reflect their knowledge of all skills?

5.6 Formative Assessment

Some of the research questions focused on the application of formative assessment. As discussed in chapter 2.5, formative assessment has proved to be an effective way to enhance learning (Black & Wiliam, 2006a, p. 9). In accordance with that, the Curriculum Guidelines for both levels advocate formative assessment (*Aðalnámskrá grunnskóla*, 2007, p. 13; *Aðalnámskrá framhaldsskóla*, 1999, p. 15). As can be seen in chapter 4.3, formative assessment seems to be more frequently practiced at the primary level than the secondary level. At both levels, written feedback on assignments is the most common form. However, this study gives no information on what types of written feedback the teachers use; whether it is what Brookhart (2007) described as effective feedback, actually giving suggestions to how to improve one’s knowledge and abilities, or whether it is merely in the form of general statements, which, according to Brookhart (2007), are not enough to be considered effective feedback (p. 56). Neither does this study reveal whether the feedback is both on grammar, on the one hand, and content and structure on the other, as Remis (2004) found that second language students in New York want (p. 82).

Some of the forms of formative assessment are, according to this study, rarely practiced in Icelandic schools. In the view of Black et al.'s (2003) argumentation that formative assessment leads to more autonomous students (pp. 66 – 67) and the emphasis laid in the Curriculum Guidelines for both levels, both regarding formative assessment and autonomous learning, it is surprising how scarce student participation in goal-setting is at both levels. According to Little (2003), one of the important aspects of learner autonomy is sharing in the goal-setting, and for students to understand where they are in their learning and where to go from there, they must be able to set goals. Thus, one would expect teachers to focus more on this aspect of autonomy, and practice formative assessment that encourages self-reflection and goal-setting.

Another form of formative assessment, which is scarcely used by the participating teachers in this study, is the formative use of summative tests. Since there is so much emphasis on written tests as this study shows, there is ample opportunity where this form is concerned. However, only four teachers at the primary level and six at the secondary level claim they use this opportunity they have of using the tests to enhance learning further. None of the teachers who teach the youngest students of 4th and 6th grades apply this form of formative assessment, though it can be said to be an ideal way to change formal tests, which the Curriculum Guidelines do not consider an appropriate way for assessment of 4th graders at least, into a method that actually improves students' learning. Hence, it is safe to say that there are substantial, unused possibilities regarding the use of written tests for the benefits of students' learning.

Finally, with 22% of the primary level and 35% of the secondary level classes deprived of formative assessment, there are strong indications here that there is a need for some re-evaluating of assessment in Icelandic schools, with the purpose of assessment kept in mind. With all the research evidence that shows formative assessment to be such a promising

way for increasing learning (Black & Wiliam, 2006a, p. 9), it can not be acceptable not to apply it. As explained in chapter 2.5, it is quite demanding for teachers to change their assessment habits, and in order for it to happen, they need to change the way they view their roles as teachers (Black et al., 2003, p. 80). Also, incentives need to be created for teachers to take part in action research or continuous education. Thus, a study of why teachers ignore formative assessment would be of interest for the educational system, as well as an extensive study, similar to that of Black et al. (2003), where teachers get assistance from specialists outside of the school as well as from each other.

5.7 Presentation of Assessment Results

The last questions, in the questionnaires, were about the presentation of final assessment results to students and, where appropriate, parents. The Curriculum Guidelines for the primary level highlight the importance of oral feedback along with the written form, which schools can have in various ways. The importance of oral feedback for the youngest students is especially emphasized (*Aðalnámskrá grunnskóla*, 2006, p. 18). However, the Curriculum Guidelines for the secondary level do not mention oral or written comments, only grades in numbers (*Aðalnámskrá framhaldsskóla*, 2004, pp. 34 – 35). One could say that there is consistency in the Curriculum Guidelines in this respect; the older the students get, the less is the need for oral feedback along with the grades given.

In accordance with this, the majority of secondary level teachers only present the grades without any comments. However, in 41% cases in this study they also give some kind of comments along with the grades, 12% use merely written comments while 29% use either oral or both written and oral comments. This shows that even though the Curriculum Guidelines do not explicitly recommend it, many teachers acknowledge the importance of oral feedback that can explain grades further. A number/letter/word grade does not say as much as

a discussion can do. With the oral feedback, teachers can explain what lies underneath the grade; which aspects are coming along well and which perhaps not so well, what the student can do to improve where improvement is needed, and so forth.

At the primary level the presentation of grades exclusively is not as common as at the secondary level. Here, grades along with written comments are the most common way of presenting the final assessment, being performed in 43% cases altogether. Oral feedback, which explicitly is recommended in the Curriculum Guidelines, is used along with the grades in only 35% cases. Where the 4th and 6th grades are concerned, however, the results are more alarming. One fourth of the participating teachers at this level only present grades, whereas only one fourth use the important oral feedback along with the grades. Since the students in the 4th and 6th grades have not come so far in their studies, the Curriculum Guidelines emphasize the importance of commenting on what they are capable of, with positive guidance to what is needed for improvement (*Aðalnámskrá grunnskóla*, 2007, pp. 17 – 18). As this is not the case, further investigation is necessary. Why do students at these beginner stages, and their parents, not get the important oral discussion on where the students are in their studies and where to go from there?

When the results about the practice of formative assessment are viewed in relation to these results, it is actually surprising that the primary school teachers do not use the oral feedback more. In 78% cases they do apply some form of formative assessment in their teaching, but when it comes to presenting the final assessment, only 35% of the classes get oral feedback with their grades.

6 Conclusion

As can be seen from this study, and the theories discussed in the thesis, language assessment can never be said to be simple and straight forward. There are various factors that need to be considered when assessment methods are chosen, whereof the purpose of the assessment is an important one. The main aim of this study was to describe how English learning is assessed in Icelandic primary and secondary schools. Studying the results, it is important to bear in mind that the questionnaires were primarily based on closed items instead of open ended ones, which can affect the results in some ways since not all people have the same understanding of words and phrases. Also, the results are based on teachers' reports on what they do, not an impartial observation. Furthermore, the sample of participants was rather small, meaning that statistical significance is not necessarily present, only indications. However, some of the indications are quite strong and need to be taken seriously.

The results show that while written tests are the leading assessment method, most teachers use various other methods and tasks as well, even though some of the important alternative assessment methods discussed in this thesis are only used by the minority of participating teachers of both levels. However, the rather infrequent and limited use of formative assessment at both levels is a matter of concern, and points to a need for re-evaluation of assessment practices in Icelandic schools. Is this because many Icelandic teachers have not examined the existing research evidence that highlights the importance of formative assessment, and the Curriculum Guidelines seem to be based on? Is this due to inadequate conditions in Icelandic schools, leaving the teachers unable to practice this form of assessment? Or is this, perhaps, due to lack of training in how to practice this effective assessment approach? These questions need to be answered in hope of bringing formative assessment and its benefits to Icelandic students of English.

The results, discussed in this thesis, have prompted many questions that are interesting subjects for further studies within the field of language assessment in Iceland. The first is the need for further investigation of written tests as discussed in chapter 5.2. Since written tests seem to be the leading method used for assessing English learning in Iceland, it is really interesting to find out whether Bachman and Palmer's philosophy of language testing (see chapter 2.4.1) is kept in mind when tests are made and implemented.

Secondly, studying why English teachers do not assess all language skills is of great interest. How can assessment results be the basis for valid inferences of language proficiency if not all language skills are included? The focus should be on what can be done to change the fact that oral and listening skills are excluded from language assessment in too many classes of various age levels around the country.

Thirdly, yet another interesting research subject is how portfolio assessment is practiced and why it is not more commonly implemented. In addition, I think it would be especially interesting to study how the time consuming factor is overcome, where this type of assessment is already practiced.

Fourthly, a study of why teachers ignore formative assessment would be of interest, as well as how it is actually implemented today by the teachers who have adopted it. As mentioned in chapter 5.6, an extensive study, similar to that of Black et al. (2003), where teachers get assistance from specialists outside of the school as well as from each other, would be worth conducting in Iceland, with the aim of enlarging the population of English teachers in Iceland who choose to apply formative assessment in their teaching and assessment processes.

Also, a study of what kind of education teachers who teach beginners have would be interesting. I think that special emphasis should be placed on whether teachers have any knowledge of the process of second language acquisition and pedagogy and special

considerations of language assessment. Whether they are familiar with informal ways of assessing language abilities would also be worth investigating.

Finally, extensive research of both teaching approaches and assessment methods, with the focus on how these relate to each other, in both primary and secondary schools would be worth while before yet another National Curriculum is written.

References

- Aðalnámskrá framhaldsskóla. Erlend tungumál* (1999). Menntamálaráðuneytið. Retrieved November 15, 2008, from <http://bella.mrn.stjr.is/utgafur/aferlendtungumal.pdf>
- Aðalnámskrá framhaldsskóla. Viðauki* (2000). Menntamálaráðuneytið. Retrieved November 15, 2008, from http://bella.mrn.stjr.is/utgafur/FramEN_DAN_STAE.pdf
- Aðalnámskrá framhaldsskóla. Almennur hluti* (2004). Menntamálaráðuneytið. Retrieved November 15, 2008, from <http://bella.mrn.stjr.is/utgafur/Mmal.pdf>
- Aðalnámskrá grunnskóla. Almennur hluti* (2006). Menntamálaráðuneytið. Retrieved November 15, 2008, from http://bella.mrn.stjr.is/utgafur/agalmennurhluti_2006.pdf
- Aðalnámskrá grunnskóla. Erlend tungumál* (2007). Menntamálaráðuneytið. Retrieved November 15, 2008, from http://bella.mrn.stjr.is/utgafur/adalnamskra_grsk_erlend_mal.pdf
- Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (part 2) [Electronic version]. *Language Teaching*, 35, 79 – 113.
- Auður Torfadóttir (2005, September 15). Er námsmat í tungumálum í takt við tímann? *Netla Vef tímarit um uppeldi og menntun*. Retrieved August 21, 2008, from <http://netla.khi.is/greinar/2005/011/index.htm>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford, England: Oxford University Press.
- Bachman, L. F. (2002a). Alternative interpretations of alternative assessments: Some validity issues in educational performance assessments [Electronic version]. *Educational Measurements: Issues and Practice*, 21(3), 5 – 18.
- Bachman, L. F. (2002b). Some reflections on task-based language performance assessment [Electronic version]. *Language Testing*, 19, 453 – 476.
- Bachman, L. F., & Cohen, A. D. (1998). Language testing – SLA interfaces: An update. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition*

- and language testing research* (pp. 1 – 31). Cambridge, England: Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford, England: Oxford University Press.
- Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests [Electronic version]. *Educational Research*, 44, 279 – 288.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, England: Open University Press.
- Black, P., & Wiliam, D. (2006a). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (pp. 9 – 25). London: Sage.
- Black, P., & Wiliam, D. (2006b). The reliability of assessments. In J. Gardner (Ed.), *Assessment and learning* (pp. 119 – 131). London: Sage.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of Assessment in Education [Electronic version]. *Assessment in Education: Principles, Policy & Practice*, 11, 7 – 26.
- Brookhart, S. M. (2007). Feedback that fits [Electronic version]. *Educational Leadership*, 65(4), 54 – 59.
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: purposes, methods, and procedures [Electronic version]. *Language Testing*, 21, 360 - 389.
- Clarke, S. (2008). *Active learning through formative assessment*. London: Hodder Education.
- Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- East, M. (2007). Bilingual dictionaries in tests of L2 writing proficiency: Do they make a difference? [Electronic version]. *Language Testing*, 24, 331 – 353.

- Erna Ingibjörg Pálsdóttir (2006). *Námsmat í höndum kennara* [Electronic version].
- Unpublished M. Ed. Thesis. Háskólinn á Akureyri.
- Gardner, J. (2006). Assessment and learning: An introduction. In J. Gardner (Ed.), *Assessment and learning* (pp. 1 – 5). London: Sage.
- Gronlund, N. E. (2003). *Assessment of student achievement* (7th ed.). Boston: Pearson Education.
- Guðfinna Gunnarsdóttir (2005). *English language assessment. A pilot study of teachers' grades and standardized test results in the 10th grade in Southern Iceland*.
- Unpublished M. Paed. Thesis. Háskóli Íslands.
- Hamp-Lyons, L. (2007). The impact of language testing practices on teaching: Ideologies and alternatives. In J. Cummins & C. Davison (Eds.) *International Handbook of English Language Teaching: Vol. 15. Springer International Handbook of Information Technology in Primary and Secondary Education* [Electronic version] (pp. 487 – 504). New York: Springer.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies [Electronic version]. *The Curriculum Journal*, 16, 207 – 223.
- Harlen, W. (2006a). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (pp. 103 – 117). London: Sage.
- Harlen, W. (2006b). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 61 – 80). London: Sage.
- Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning [Electronic version]. *Assessment in Education: Principles, Policy & Practice*, 10, 169 – 207.
- James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 47 – 60). London: Sage.
- James, M., & Pedder, D. (2006). Professional learning as a condition for assessment for

- learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 27 – 43). London: Sage.
- Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: assessment as discourse and assessment of discourse [Electronic version]. *Language Testing*, 21, 335 – 359.
- Lilja Jóhannsdóttir (2008). *Námsmat í ensku í 8. bekk*. Unpublished M. Ed. Thesis. Kennaraháskóli Íslands.
- Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching* (9th ed.). Upper Saddle River, NJ: Pearson Education.
- Little, D. (2003, January 16). *Learner autonomy and second/foreign language learning*. Retrieved October 21, 2008, from <http://www.llas.ac.uk/resources/gpg/1409>
- Manitoba Education, Citizenship and Youth (2006). *Rethinking classroom assessment with purpose in mind: assessment for learning, assessment as learning, assessment of learning*. Manitoba, Canada: Author. Retrieved January 19, 2009 from <http://www.edu.gov.mb.ca/k12/assess/wncp/index.html>
- McKay, P. (2006). *Assessing young language learners*. Cambridge, England: Cambridge University Press.
- Menntamálaráðuneytið (n.d.a). <http://www.menntamalaraduneyti.is/stofnanir>
- Menntamálaráðuneytið (n.d.b). <http://www.menntamalaraduneyti.is/utgefid-efni/namskrar//nr/3953>
- Menntamálaráðuneytið (n.d.c). <http://www.menntamalaraduneyti.is/utgefid-efni/namskrar//nr/3954>
- Nunan, D. (1992). *Research methods in language learning*. Cambridge, England: Cambridge University Press.
- Ragnheiður Hermannsdóttir (2008). *Það skiptir svo miklu máli hvernig þetta er gert fyrir námið: námsmat frá sjónarhóli nemenda* [Electronic version]. Unpublished M. Ed. Thesis. Kennaraháskóli Íslands.

- Rea-Dickins, P. (2007). Classroom-based assessment: Possibilities and pitfalls. In J. Cummins & C. Davison (Eds.) *International Handbook of English Language Teaching: Vol. 15. Springer International Handbook of Information Technology in Primary and Secondary Education* [Electronic version] (pp. 505 – 520). New York: Springer.
- Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: disentangling the construct of formative assessment [Electronic version]. *Language Testing*, 17, 215 – 243.
- Remis, Milena (2004). *Students and feedback: A survey of students' views on feedback in second language writing*. Unpublished M. Paed. Thesis. Háskóli Íslands.
- Rósa Maggý Grétarsdóttir (2007). Vandinn að velja – um námsmatsaðferðir og tilgang námsmats [Electronic version]. *Gátt. Ársrit um fullorðinsfræðslu og starfsmenntun*, 35 – 39.
- Rúnar Sigþórsson (2008). *Mat í þágu náms eða nám í þágu mats: samræmd próf, kennsluhugmyndir kennara, kennsla* [Electronic version]. Reykjavík, Iceland: Kennaraháskóli Íslands.
- Saito, H. (2008). EFL classroom peer assessment: Training effects on rating and commenting [Electronic version]. *Language Testing*, 25, 553 – 581.
- Shepard, L. A. (2000). The role of assessment in a learning culture [Electronic version]. *Educational Researcher*, 29(7), 4 – 14.
- Steinlaug Sigríður Bjarnadóttir (2008). *Written feedback in foreign English writing instruction. A study of teacher feedback in the 8th and 10th grades of the compulsory school in Iceland*. Unpublished M. Paed. Thesis. Háskóli Ísland.
- Stiggins, R. (2004 September). New assessment beliefs for a new school mission. *Phi Delta Kappan* 86(1). Retrieved November 20, 2008, from <http://proquest.umi.com/pqdlink?index=7&did=694211911&SrchMode=3&sid=2&Fmt=3&VInst=PROD&VType=PQD&RQT=309&VName=PQD&TS=1227177630&clientId=58032&aid=1>

- Stiggins, R. (2007). Assessment through the student's eyes [Electronic version]. *Educational Leadership*, 64(8), 22 – 26.
- Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment: What's in it for schools?* London: Routledge Falmer.
- William, D. (2007). Changing classroom practice [Electronic version]. *Educational Leadership*, 65(4), 36 – 42.
- Zangl, R. (2000). Monitoring language skills in Austrian primary (elementary) schools: case study [Electronic version]. *Language Testing*, 17, 250 – 260.

Appendix

A Questionnaire for Primary Level Teachers – English Version

1. Information on the school

1.1 How big is the school?

- ☐ Fewer than 100 students ☐ 100 – 400 students ☐ Over 400 students

1.2 Which grade level is this?

- ☐ 4th grade ☐ 6th grade ☐ 8th grade ☐ 10th grade

2. Information on the teacher

2.1 What is your education?

- ☐ B. Ed. degree, majoring in English
- ☐ B. Ed. degree, not majoring in English
- ☐ B. A. (or a higher degree) in English with teaching certificate
- ☐ B. A (or a higher degree) in English without teaching certificate
- ☐ Other

2.2 How long have you been teaching English?

- ☐ Less than 5 years ☐ 5 – 10 years ☐ 11 – 20 years ☐ Over 20 years

3. Assessment methods. Here you will be asked what assessment methods are used to assign the final grade in your year.

3.1 Do you use written tests?

- ☐ Yes ☐ No

3.2 If yes to question 3.1, what is the total percentage of the final grade?

3.3 If yes to question 3.1, what language skills are assessed that way? Please tick all relevant skills.

- | | | |
|-------------------------------------|----------------------------------|-----------------------------------|
| <input type="checkbox"/> Listening | <input type="checkbox"/> Reading | <input type="checkbox"/> Writing |
| <input type="checkbox"/> Vocabulary | <input type="checkbox"/> Grammar | <input type="checkbox"/> Spelling |

3.4 If yes to question 3.1, what types of tasks are used? Please tick all relevant aspects.

- | | | |
|---|---|---|
| <input type="checkbox"/> Multiple-choice items | <input type="checkbox"/> Open questions | <input type="checkbox"/> Gap-filling items |
| <input type="checkbox"/> True/false items | <input type="checkbox"/> Matching items | <input type="checkbox"/> Translation into Icelandic |
| <input type="checkbox"/> Pre-read texts | <input type="checkbox"/> Unseen texts | <input type="checkbox"/> Free writing |
| <input type="checkbox"/> Translation into English | | <input type="checkbox"/> Other |

3.5 If 'other' in question 3.4, please explain here:

3.6 If yes to question 3.1, are there time limits?

- ☐ Yes ☐ No

3.7 If yes to question 3.6, why and how is it decided?

3.8 If yes to question 3.1, what types of tests are used? Please tick all relevant boxes.

- ☐ Traditional tests
- ☐ Tests allowing some means of learning aids, such as dictionaries
- ☐ Cooperation tests
- ☐ Tests taken in more than one day, allowing students to take questions home to study between days.
- ☐ Take-home tests
- ☐ Other

3.9 If 'other' in question 3.8, please explain here:

3.10 Do you use written assignments?

☐ Yes ☐ No

3.11 If yes to question 3.10, what is the total percentage of the final grade?

3.12 If yes to question 3.10, what language skills are assessed that way? Please tick all relevant skills.

☐ Listening ☐ Reading ☐ Writing
☐ Vocabulary ☐ Grammar ☐ Spelling

3.13 If yes to question 3.10, what types of tasks are used? Please tick all relevant boxes.

☐ Retelling ☐ Summaries ☐ Book reviews ☐ Essays
☐ Free writing ☐ Translation into English ☐ Other

3.14 If 'other' in question 3.13, please explain here:

3.15 Do you use oral tests?

☐ Yes ☐ No

3.16 If yes to question 3.15, what is the total percentage of the final grade?

3.17 If yes to question 3.15, what language skills are assessed that way? Please tick all relevant skills.

☐ Listening ☐ Speaking ☐ Reading
☐ Vocabulary ☐ Grammar

3.18 If yes to question 3.15, what types of tasks are used? Please tick all relevant boxes.

☐ Pre-read text discussed ☐ Previously viewed film/TV-show discussed
☐ Pictures viewed and discussed ☐ Discussion on daily life
☐ Discussion between students ☐ Other

3.19 If 'other' in question 3.18, please explain here:

3.20 Do you use oral presentations?

☐ Yes ☐ No

3.21 If yes to question 3.20, what is the total percentage of the final grade?

3.22 If yes to question 3.20, what language skills are assessed that way? Please tick all relevant skills.

☐ Listening ☐ Speaking ☐ Reading ☐ Writing

☐ Vocabulary ☐ Grammar ☐ Spelling

3.23 If yes to question 3.20, what types of tasks are used? Please tick all relevant aspects.

☐ Presentation of pre-read material ☐ Presentation of film or TV-show

☐ Presentation of own choice ☐ Power point presentation ☐ Other

3.24 If 'other' in question 3.23, please explain here:

3.25 Do you use self-assessment?

☐ Yes ☐ No

3.26 If yes to question 3.25, what is the total percentage of the final grade?

3.27 If yes to question 3.25, what language skills are assessed that way? Please tick all relevant skills.

☐ Listening ☐ Speaking ☐ Reading ☐ Writing

☐ Vocabulary ☐ Grammar ☐ Spelling

3.28 Do you use peer-assessment?

☐ Yes ☐ No

3.29 If yes to question 3.28, what is the total percentage of the final grade?

3.30 If yes to question 3.28, what language skills are assessed that way? Please tick all relevant skills.

☐ Listening ☐ Speaking ☐ Reading ☐ Writing

☐ Vocabulary ☐ Grammar ☐ Spelling

3.31 Do you use portfolio assessment?

☐ Yes ☐ No

3.32 If yes to question 3.31, what is the total percentage of the final grade?

3.33 If yes to question 3.31, what language skills are assessed that way? Please tick all relevant skills.

☐ Listening ☐ Speaking ☐ Reading ☐ Writing

☐ Vocabulary ☐ Grammar ☐ Spelling

3.34 If yes to question 3.31, what type of portfolio assessment? Please tick all relevant boxes.

☐ Logbooks ☐ Process portfolio

☐ Showcase portfolio ☐ Other

3.35 If 'other' in question 3.34, please explain here:

3.36 Are there any other methods used or factors included in the final grade? Please explain here:

4. Formative Assessment

4.1 Is formative assessment practiced?

☐ Yes ☐ No

4.2 If yes to question 4.1, what form of formative assessment? Please tick all relevant boxes.

- | | |
|---|--|
| <input type="checkbox"/> Written feedback on assignments | <input type="checkbox"/> Oral feedback on assignments |
| <input type="checkbox"/> Oral feedback in discussion | <input type="checkbox"/> Teacher-student interviews |
| <input type="checkbox"/> Interviews in small groups | <input type="checkbox"/> Student participation in goal-setting |
| <input type="checkbox"/> Formative use of summative tests | <input type="checkbox"/> Other |

4.3 If 'formative use of summative tests' in question 4.2, please explain here:

4.4 If 'other' in question 4.2, please explain here:

5. Presentation of assessment results.

5.1 How are final assessment results presented to students and, where appropriate, parents?

- ☐ Grades in numbers/letters/words
- ☐ Written comments
- ☐ Grades in numbers/letters/words along with written comments
- ☐ Grades in numbers/letters/words along with oral comments
- ☐ Grades in numbers/letters/words along with written and oral comments

5.2 Is there anything else you would like to mention regarding presentation of assessment results?

A Questionnaire for Primary Level Teachers – Icelandic Version

Könnun á námsmatsaðferðum í ensku í grunnskólum á Íslandi

1. Grunnupplýsingar um skólann

1.1 Hversu stór er skólinn?

☐ færri en 100 nemendur ☐ 100 - 400 nemendur ☐ fleiri en 400 nemendur

* 1.2 Hvaða árgang er um að ræða hér?

☐ 4. bekk ☐ 6. bekk ☐ 8. bekk ☐ 10. bekk

2. Grunnupplýsingar um kennarann

2.1 Hvaða menntun hefur þú?

☐ B.Ed próf með ensku sem aðalfag ☐ B.Ed. próf ekki með ensku sem aðalfag ☐ B.A próf (eða hærri prófgráðu) í ensku ásamt kennsluréttindum ☐ B.A. próf (eða hærri prófgráðu) í ensku án kennsluréttinda ☐ Annað

2.2 Hve lengi hefur þú kennt ensku?

☐ Skemur en 5 ár ☐ 5 til 10 ár ☐ 11 til 20 ár ☐ Lengur en 20 ár

3. Matsaðferðir

Undir þessum lið er spurt um matsaðferðir og þætti sem lokamat byggir á.

3.1 Eru notuð skrifleg próf?

☐ Já ☐ Nei

3.2 Ef já við spurningu 3.1: Hvert er samtals hlutfall af lokaeinkunn?

3.3 Ef já við spurningu 3.1: Hvaða þættir eru metnir þannig?

Vinsamlegast merkið við alla þætti sem við eiga.

☐ Hlustun ☐ Lesskilningur ☐ Ritun ☐ Orðaforði ☐ Málfræði ☐ Stafsetning

3.4 Ef já við spurningu 3.1: Á hvernig verkefnum byggja prófin?

Vinsamlegast merkið við alla þætti sem við eiga.

☐ Krossaspurningum ☐ Opnum spurningum ☐ Eyðufyllingum ☐ Rétt/rangt atriðum
☐ Pörun atriða ☐ Þýðingum ☐ Lesnum textum ☐ Ólesnum textum ☐ Frjálsri ritun
☐ Stílum ☐ Annað

3.5 Ef krossað var við 'annað' í spurningu 3.4: Vinsamlegast útskýrið hér

3.6 Ef já við spurningu 3.1: Eru tímamörk?

☐ Já ☐ Nei

3.7 Ef já við spurningu 3.6: Hvers vegna og hvernig eru tímamörk ákveðin?

3.8 Ef já við spurningu 3.1: Hvernig gerðir af skriflegum prófum?

Vinsamlegast merkið við alla þætti sem við eiga

☐ Hefðbundin skrifleg próf ☐ Gagnapróf (þ.e. einhvers konar hjálpargögn leyfð)

☐ Samvinnupróf ☐ Próf sem tekin eru á fleiri dögum og nem. geta farið með spurningar heim milli daga til að læra betur ☐ Heimapróf ☐ Annað

3.9 Ef krossað er við 'annað' í spurningu 3.8: Vinsamlegast útskýrið hér

3.10 Eru notuð ritunarverkefni?

☐ Já ☐ Nei

3.11 Ef já við spurningu 3.10: Hvert er samtals hlutfall af lokaeinkunn?

3.12 Ef já við spurningu 3.10: Hvaða þættir eru metnir þannig?

Vinsamlegast merkið við alla þætti sem við eiga.

☐ Hlustun ☐ Lesskilningur ☐ Ritun ☐ Orðaforði ☐ Málfræði ☐ Stafsetning

3.13 Ef já við spurningu 3.10: Hvernig verkefni er um að ræða?

Vinsamlegast merkið við alla þætti sem við eiga.

☐ Endursagnir ☐ Útdrættir ☐ Bókarýni ☐ Ritgerðir ☐ Frjáls ritun ☐ Stílar ☐ Annað

3.14 Ef krossað var við 'annað' í spurningu 3.13: Vinsamlegast útskýrið hér

3.15 Eru notuð munnleg próf?

☐ Já ☐ Nei

3.16 Ef já við spurningu 3.15: Hvert er samtals hlutfall af lokaeinkunn?

3.17 Ef já við spurningu 3.15: Hvaða þættir eru metnir þannig?

Vinsamlegast merkið við alla þætti sem við eiga.

☐ Hlustun ☐ Talað mál ☐ Lesskilningur ☐ Orðaforði ☐ Málfræði

3.18 Ef já við spurningu 3.15: Á hvernig verkefnum byggja prófin?

Vinsamlegast merkið við alla þætti sem við eiga.

- ☐ Fyrirfram lesinn texti til umfjöllunar ☐ Fyrirfram skoðuð kvikmynd/sjónvarpsþáttur til umfjöllunar ☐ Myndir, skoðaðar á staðnum, til umfjöllunar ☐ Umfjöllunarefni úr daglega lífinu ☐ Samtöl milli nemenda ☐ Annað

3.19 Ef krossað var við 'annað' í spurningu 3.18: Vinsamlegast útskýrið hér

3.20 Eru notaðar munnlegar kynningar á verkefnum?

- ☐ Já ☐ Nei

3.21 Ef já við spurningu 3.20: Hvert er samtals hlutfall af lokaeinkunn?

3.22 Ef já við spurningu 3.20: Hvaða þættir eru metnir þannig?

Vinsamlegast merkið við alla þætti sem við eiga.

- ☐ Hlustun ☐ Talað mál ☐ Lesskilningur ☐ Ritun ☐ Orðaforði ☐ Málfræði ☐ Stafsetning

3.23 Ef já við spurningu 3.20: Hvers konar verkefni er um að ræða?

Vinsamlegast merkið við alla þætti sem við eiga.

- ☐ Kynning á lesnu efni ☐ Kynning á kvikmynd/sjónvarpsþætti ☐ Munnleg umfjöllun út frá eigin brjósti ☐ Glærukynning ☐ Annað

3.24 Ef krossað var við 'annað' í spurningu 3.23: Vinsamlegast útskýrið hér

3.25 Er notað sjálfsmat?

- ☐ Já ☐ Nei

3.26 Ef já við spurningu 3.25: Hvert er samtals hlutfall af lokaeinkunn?

3.27 Ef já við spurningu 3.25: Hvaða þættir eru metnir þannig?

Vinsamlegast merkið við alla þætti sem við eiga.

- ☐ Hlustun ☐ Talað mál ☐ Lesskilningur ☐ Ritun ☐ Orðaforði ☐ Málfræði ☐ Stafsetning

3.28 Er notað jafningjamat?

- ☐ Já ☐ Nei

3.29 Ef já við spurningu 3.28: Hvert er samtals hlutfall af lokaekunn?

3.30 Ef já við spurningu 3.28: Hvaða þættir eru metnir þannig?

Vinsamlegast merkið við alla þætti sem við eiga.

- ☐ Hlustun ☐ Talað mál ☐ Lesskilningur ☐ Ritun ☐ Orðaforði ☐ Málfræði
☐ Stafsetning

3.31 Er notað 'portfolio'-mat?

- ☐ Já ☐ Nei

3.32 Ef já við spurningu 3.31: Hvert er samtals hlutfall af lokaekunn?

3.33 Ef já við spurningu 3.31: Hvaða þættir eru metnir þannig?

Vinsamlegast merkið við alla þætti sem við eiga.

- ☐ Hlustun ☐ Talað mál ☐ Lesskilningur ☐ Ritun ☐ Orðaforði ☐ Málfræði
☐ Stafsetning

3.34 Ef já við spurningu 3.31: Hvaða form er á matinu?

- ☐ Leiðarbækur ☐ Safnmöppur ☐ Sýnismöppur ☐ Annað

3.35 Ef krossað er við 'annað' í spurningu 3.34; Vinsamlegast útskýrið hér

3.36 Eru aðrar matsaðferðir eða -þættir en nefndir hafa verið hér að ofan notaðir? Vinsamlegast útskýrið.

4. Leiðsagnarmat

4.1 Er leiðsagnarmati beitt?

- ☐ Já ☐ Nei

4.2 Ef já við spurningu 4.1: Í hvaða formi er það?

Vinsamlegast merkið við alla þætti sem við eiga.

- ☐ Skrifleg endurgjöf við verkefni ☐ Munnleg endurgjöf við verkefni ☐ Munnleg endurgjöf í umræðum ☐ Einstaklingsviðtöl við nemendur ☐ Viðtöl við nemendur í litlum hópum
☐ Þátttaka nemenda í markmiðssetningu ☐ Unnið með niðurstöður prófa þannig að nem. hafi tækifæri á að bæta sig ☐ Annað

4.3 Ef krossað var við 'unnið með niðurstöður prófa' í spurningu 4.2: Vinsamlegast útskýrið hér

4.4 Ef krossað var við 'annað' í spurningu 4.2: Vinsamlegast útskýrið hér

5. Kynning á lokamati

5.1 Hvernig er lokamat kynnt nemendum og foreldrum?

- ☐ Eingöngu einkunn í formi tölustafa/bókstafa/orða
 ☐ Eingöngu skrifleg umsögn
 ☐ Einkunn í formi tölustafa/bókstafa/orða með skriflegri umsögn
 ☐ Einkunn í formi tölustafa/bókstafa/orða með munnlegri umsögn
 ☐ Einkunn í formi tölustafa/bókstafa/orða með skriflegri og munnlegri umsögn

5.2 Er eitthvað annað sem þú vilt koma á framfæri varðandi kynningu á lokamati?

A Questionnaire for Secondary Level Teachers – English Version

1. Information on the school

1.1 Which course unit is this?

☐ ENS 202 ☐ ENS 203 ☐ ENS 303 ☐ ENS 403 ☐ ENS 503

2. Information on the teacher

2.1 What is your education?

- ☐ B. A. degree in English with teaching certificate
- ☐ B. A. degree in English without teaching certificate
- ☐ Masters (or a higher) degree in English with teaching certificate
- ☐ Masters (or a higher) degree in English without teaching certificate
- ☐ Other

Other questions were the same in both questionnaires.

A Questionnaire for Secondary Level Teachers – Icelandic Version

Könnun á námsmatsaðferðum í ensku í framhaldsskólum á Íslandi

1. Grunnupplýsingar um skólann

* 1.1 Hvaða áfanga er um að ræða hér?

☐ ENS 202 ☐ ENS 203 ☐ ENS 303 ☐ ENS 403 ☐ ENS 503

2. Grunnupplýsingar um kennarann

2.1 Hvaða menntun hefur þú?

- ☐ B.A. próf í ensku ásamt kennsluréttindum ☐ B.A. próf í ensku án kennsluréttinda
- ☐ Masterspróf (eða hærri prófgráðu) í ensku ásamt kennsluréttindum ☐ Masterspróf (eða hærri prófgráðu) í ensku án kennsluréttinda ☐ Annað