



Title	The effects of assessment on stakeholders within an elementary-grade EFL program
Author(s)	Crosthwaite, PR
Citation	Asian EFL Journal, 2009
Issued Date	2009
URL	http://hdl.handle.net/10722/205826
Rights	Creative Commons: Attribution 3.0 Hong Kong License

The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program.

Peter Robert Crosthwaite.

Abstract

This paper will explore which aspects of English language assessment could be said to have affected the teaching and learning that took place within the context of an elementary grade EFL program, from the point of view of the different stakeholders involved. This paper presents, by way of a case study, an exploration of the effects of summative, formative and criterion-based aspects of assessment and the impact that they had on the stakeholders in the EFL program. The case study explores attitudes towards the introduction of explicitly defined learning outcomes in the EFL program's curriculum to see how they affected approaches and attitudes towards learning and teaching. The second area to be explored was how the inclusion of monthly spoken language tests affected the test stakeholders' perceptions of learning and teaching English as a foreign language.

Background and Rationale

In the year 2007 I was working in the foreign language institute of a national university in South Korea. This institute provides a wide range of EFL programs, the most popular of which is an elementary school-grade general English education program, with around 250 students in total aged from seven to thirteen years old and with a staff of 6 native Korean English teachers and 6 native English-speaking teachers.

During my tenure at the institute during 2007, parents often criticized the institute's
Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

EFL program for a lack of focus, citing a lack of specific aims or objectives or outcomes. There wasn't any planned curriculum in operation for the EFL program other than the use of a variety of popular EFL textbooks allocated to us from the language institute's administrative office. In terms of a syllabus, teachers were encouraged by the director of studies to divide the number of classes over a three month period by the number of pages in the assigned textbook in order to see how many pages they would teach per class. Teachers would then enter the page numbers they were going to teach over the three month period into a word-processed document that served as the *de facto* 'curriculum' for the program. This kind of approach towards curriculum making was one that I had encountered throughout my entire work experience at the time.

As the curriculum consisted merely of page numbers, communication about the students' progress on the curriculum between fellow teachers was almost non-existent, which meant that little to no co-operative lesson planning took place. As the classes were shared between a native speaker and a Korean English teacher, this lack of communication became a serious problem as it meant that the classes became un-coordinated, leading to confusion from students and parents and teachers about 'whom' should be teaching 'what'. The KETs were supposed to give the students lessons on the target forms in Korean, drilling them on vocabulary and grammar, as well as their written English, and the NETs were supposed to follow up the Korean English lesson by giving the students the opportunity to practice outputting the forms they had been drilled on. However, if the students had difficulty in producing meaningful output of the target forms to the NET, this situation was not often correctly communicated to the KET who would simply move on with the content of the book, leaving those students behind, rather than providing any

opportunities for review that might have helped the students to correct their mistakes. This usually led to wide differences in the ability of the students in a particular class if put together over an extended period of time.

In terms of assessment, as there wasn't a standard curriculum to follow whereby students' performance could be properly assessed once they had entered the program, this in effect meant that it was not possible for teachers to create much of a framework for useful internal assessment. The only internal assessment that was done during the program was a bi-monthly written report to parents done by the native English speakers, which was therefore written in English. Parents could not easily understand the comments made by the native English speaking teachers and therefore the assessments were considered a waste of time, yet were still required paperwork according to the administrative office. In order to avoid complaints, teachers often wrote glowing reports in praise of each student, with negative comments being re-written at the institute's administrative office's request before being delivered to parents. Over the year however, those parents who had been able to read the reports often questioned why the reports looked so good despite an apparent lack of progress from their children. In short, the old program could be said to use summative assessment of the worst kind – a fabricated response dressed up to save face for teachers and parents, serving no true summative (or formative) function whatsoever.

For these reasons, I was asked to develop a new curriculum with new assessment procedures by January 2008 in an attempt to relieve the problems faced by the stakeholders at the institute. I will list the main provisions in the new curriculum as follows:

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

1) This new curriculum would involve a set of weekly ‘learning objectives’ and these objectives were to be written in English (unfortunately due to time constraints they would not be translated into Korean, which I felt was going to be a big problem from the offset). The objectives were to be based on the content of a new set of textbooks (Macmillan’s ‘Treasures’ series) that we had ordered for the new school year. Students, parents, and teachers would be given a copy of these objectives for the month, making the monthly lesson plans explicit to all stakeholders involved. It was hoped that the tasks would allow the students to ‘focus on form’ while performing the tasks specified in the syllabus (Long, 1985). To explain further, these objectives were intended to be what Long defines as ‘target tasks’ rather than ‘pedagogic’ tasks. Pedagogic tasks (which would be the ‘how’ of teaching) would be devised by the teachers themselves as long as these tasks helped students to the outcome specified in the learning objective or ‘target task’. These target tasks could be said to be part of a rhetorical classification of tasks, defined by Swales (1990) as a ‘class of communicative events, the members of which share some set of communicative purposes.’ (Swales, 1990:58). The learning objectives were intended to be examples of such ‘communicative purposes’, the thematic content of which were to be formulated from the various discourse domains encountered in the textbooks. A sample of these objectives/tasks can be found in appendix 1. However, as the content of the tasks had to follow the sequencing of the content of the textbook, the choice of task types was limited, and we could not pay proper attention to the correct sequencing of tasks according to complexity, which is Ellis’ (2003) fourth stage in his guide to constructing a useful task based curriculum. It was hoped that the process of following the tasks would be enough for the benefits of TBLT (that of increased implicit and explicit knowledge of language

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

through tasks with a focus on form) to be felt in the classroom.

2) There would now be a speaking test performed every month in order to see if the students were meeting the learning objectives laid out in the curriculum. These tests were comprised of 10 questions based on the grammar forms, vocabulary and dialogs found in that classes' assigned textbook with an aim to emulate the 'can-do' function of the objectives. The tests would be performed by teachers one-on-one with the students towards the end of each month. Teachers could not test their own classes' students and had to test the classes of other teachers, in a bid to avoid fixing their own classes' scores. Therefore, each teacher had to write their own test questions so that the exact wording of each question would remain secret until the test. The tests were to be graded on a scale of 0-10, with 10 being the highest mark available and 0 being the lowest. The grading scale was as follows:

- 10 – full sentence(s) given without hesitation with perfect pronunciation
- 9- full sentence(s) given without hesitation with good pronunciation
- 8 – full sentence(s) given with some small hesitation with good pronunciation
- 7- full sentence(s) given with some hesitation and/or poor pronunciation
- 6- attempt at full sentence(s) made with good pronunciation
- 5 – attempt at full sentence(s) made, poorly pronounced
- 4- one word answer, good pronunciation, no hesitation
- 3 – one word answer, good pronunciation, some hesitation
- 2 – incorrect answer, with good pronunciation, no hesitation
- 1 – incorrect answer, with hesitation and or poor pronunciation
- 0 – did not answer the question

These grading scales were formulated through discussions between me and my co-teachers, but with no real sound reference to any established grade scales. This is

problematic in that the validity of the grades could easily be called into question, particularly as the native speakers actually giving the tests may have differed in their interpretation of a student's performance on this grading scale. Godfrey and Hoekje (1990) questioned the viability of using native speaker (NS) norms to evaluate nonnative speaker texts in the absence of concrete data on NS performance (in Hoekje and Linnel, 1994:109), and while their research was targeting non-native assistant English teachers' performances on speaking tests, the same conclusions could possibly be drawn in this context also – particularly in the area of pronunciation. It was possible that whatever constituted 'perfect' and 'good' pronunciation for grade scale markers 9 to 10 could be confused, or that the criteria for 'hesitation' could be interpreted differently by different NSs in the absence of clear established guidelines for grading these section.

However, as the test questions were formulated from the tasks in the curriculum, we hoped that the tests would have the effect found in Brindley's (1994) findings on 'task-centered assessment' – that of 'both teachers and learners focusing on language as a tool' (Brindley, 1994, in Ellis, 2003:279). It was also hoped that the tests would achieve a positive washback effect, with the results of the assessment to be integrated into the learning process – thus following the criteria for formative assessment. Whether the tasks achieved these aims or not will be discussed in the commentaries of the case studies and the subsequent discussion.

3) The results of these tests as a percentage were given to parents in the form of a report, with attached comments from the teachers regarding the progress and attitude of the learners. This report contained the test questions that the students had answered, as well

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

as their performances on the grading scale given in 2) of this section (to see a sample form of the test questions and report, please see appendix 2). Every three months, students and parents would receive their average grades over this period, as well as the average three-month average grade of the class, for parents and students to compare their performance against other students in the class.

4) Students who had gained an average score for three months of tests of under 60% would have to be level-tested under the initial level test conditions once again, and there was a possibility that they may have had to enter a lower level class.

5) As each level of the program (there are 6 distinct levels, separated by the levels that the textbooks follow) took one year to complete, students who gained an average 60% of their grades over one year would get a 'pass' grade, and move up to the next level at a graduation ceremony at the end of the academic year. Students who received an average grade of under 60% for the year would have to repeat the level again the following year.

The idea behind the new curriculum and guidelines from my seniors in charge of administration was that the new, more specific criteria for assessment would promote teamwork between Korean teachers, native English speaking teachers, and parents, as well as to provide some measure of students' relative successes or failures during the program. I devised the 'learning objectives' for the year for each level of the program, and the program was implemented in March 2008, in time for the next academic year.

To summarize the main points of the implementations, the program had:

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

- 1) Introduced an objective-based curriculum, with the criterion for these objectives formed from the content of the Macmillan EFL textbook series titled ‘Treasures’.
- 2) Created a system where students may or may not complete the goals defined in the criteria of the objective-based curriculum (known as criteria-based assessment), with completion of said criteria measured by introducing 3)
- 3) A monthly testing procedure where percentage grades were calculated as an average percentage over a three-month period, with students comparing their three-month grade averages with other classmates and their parents. Students were to be re-level-tested if an average score of 60% was not reached over the three month period.

Opposing elements within assessment and their possible impact on EFL contexts, a review.

The role of assessment in the classroom is of high importance. Assessment is the ongoing process where what is learned and taught in the classroom is tested and recorded over time. Assessment is undertaken for different purposes depending on the nature of the context the assessment will be used in, including formative and summative assessment, as well as criterion referenced and norm referenced assessment. However, when we think about classroom assessment, how can we be sure that the kind of assessment used in a specific context constitutes good practice? How can we be sure that what is planned, taught, tested and learned from assessments actually benefits the learners in any way? It is therefore necessary to examine and criticize some of the different approaches within assessment by reviewing some of the opinions laid out by researchers who are involved in the assessment

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

field, in a bid to give some background into the case study this paper will later present. For reference, a brief overview of four different approaches to assessment is given below.

Formative and Summative Assessment

Assessment can usually be broken down into two categories that highlight the general purpose of the assessment taking place:

Summative assessment is generally meant to measure the achievement of learners over a given period of time, such as at the end of a course module, and it is also used for standardized tests where a grade or score needs to be given in order to prove a learner's ability. Scores and grades from these tests usually have very high stakes for those involved, and are used to compare the quality of schools, language programs, job applicants, quality of university entrants and a wide range of other categories. For summative assessment, 'the focus of attention is on the result of learning [and is] concerned with identifying overall levels of achievement and measuring what learners do against them.' (Rea-Dickens, 2000:377). This kind of assessment is usually called 'assessment *of* learning', and is usually not intended to guide the learning process in any way, but merely to measure the learning that has taken place.

Formative assessment looks at the learning process itself in a bid to guide the learning that takes place. It is usually conducted by classroom teachers in relation to what is taught in class in order to diagnose students' knowledge and abilities and to act as a guide of what to next do to improve them. This kind of assessment is usually called 'assessment *for* learning', where the assessment spots areas of trouble and is used to guide

future classes, feedback, or revision of the curriculum. The washback received from the assessments is the fuel that will power the next focus of classroom practice.

Criterion and Norm-reference based assessment.

When deciding where the questions asked on a test usually come from, and for determining whether students fail or succeed the assessment that is to take place, assessment is normally based on either a criterion-based or norm-referenced approach.

Criterion-based assessment involves the use of pre-defined learning outcomes or goals that are set in the curriculum or in the grading rubrics of standardized tests. These goals are set by the curriculum makers, usually either the director of studies in an EFL/ESL language program, or as is often the case, found within the teacher manuals of most western EFL/ESL textbook programs, which the teachers then use to create their in-house tests for assessment purposes.

Norm-referenced assessment does not have pre-defined learning goals but instead uses a curve-type grading system to divide the grades between the students in a particular class or category. It compares the results of all the students in the class and adjusts the grades accordingly according to the grading requirements put in place by an institution or the performance of the students as an entire class. This means that these kinds of grades are awarded to students often regardless of whether the students could successfully complete the tasks they have been given to do during the assessment period, as long as their test score fell within the boundary created by the grading curve. As the grades within the reports

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

given to parents in the past were always high, regardless of actual assessment performance, it could be said that the EFL program in my context followed this form of assessment, and it was my goal to move away from this approach in our context.

Issues within assessment – what constitutes good practice?

It would now be useful to take on the opinions of researchers who work in the field of assessment in order to see which forms of assessment could be said to constitute good practice. Sadler (2007) sets the general assessment scene used in most countries worldwide today as having:

‘high levels of external societal demand for qualifications and certificates, adoption of criterion-based marking and assessment, and the commitment of teachers to do their very best for students.’ (Sadler 2007: 388-389)

In other words, the current trend for assessment promotes a mix of formative and summative assessment, as well as the adoption of criterion-based standards against which these assessments would be based on, as shown in Sadler’s (2005) survey of 65 universities in Australia, Canada, New Zealand, South Africa, Sweden, United Kingdom, and the USA, with ‘many institutions employing identical or related models without necessarily calling them criteria-based.’ (Sadler, 2005:176).

However, Sadler goes on to mention that:

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

‘[...] the real state of affairs is, in many supposedly ‘learning’ situations, a long way from this. A problematic mix of mere compliance, and of going ever-so-systematically through all the steps, actually turns out to be instrumental in subverting the goal of assessment.’ (Sadler, 2007:389)

Sadler’s view is supported in Harlen and Crick (2003) who suggest that ‘...current testing practices are detrimental to, rather than encouraging of, the attitudes and energy for learning needed for lifelong learning’ (Harlen and Crick, 2003:204). Therefore, there seems to be a huge gap between the supposed vision of what constitutes good assessment practice, and the reality of the situation found in different assessment contexts.

The Dangers of Summative Assessment.

Often, the stakes of summative assessment, especially from external assessing sources, are incredibly high and have a huge impact on our lives and futures within modern society. The qualifications gained from testing can often define our whole character to others around us. One would think then, that high-stakes tests would provide all the motivation for the learner that is necessary for learning to take place. However, according to studies carried out by McNeil and Valenzuela (1998), reviewed in Harlen and Crick (2003), high stakes tests ‘do not have a markedly positive effect on teaching and learning in the classroom. High stakes tests do not motivate the unmotivated’ (Harlen and Crick, 2003:173)

There is also a huge amount of pressure on teachers and students to pass these tests, Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

a trend reviewed in Harlen and Crick (2003) who find that high stakes tests increase test anxiety (Benmansour, 1999; Leonard & Davey, 2001; Pollard *et al.*, 2000), increase the pressure on students to do well (Davies & Brember, 1998; Leonard & Davey, 2001), focus teaching on the content of the tests (Johnston & Mc-Clune, 2000) and leads to students adopting test-taking strategies that are harmful to learning development (Paris *et al.*, 1991; Reay & Wiliam, 1999)

Another key part of summative assessment is grading. Broadfoot *et al* (1992) mentions that in terms of assessment which leads to the labeling of a student's abilities with a grade, usually denoted by the letters A to F, and in our context on a scale of 0-10, this form of assessment can 'in certain circumstances have harmful effects on children's learning'(Broadfoot *et al*, 1992:6). He mentions further:

There is much research... that shows that a competitive climate, such as labelling produces, does not increase learning outcomes for all. Rather the already successful thrive and the less successful (always the majority) underachieve more and more as they are repeatedly discouraged by the way their efforts are judged (Broadfoot *et al*, 1992:6).

This view is supported in Torrance (2007), who claims that achievements in summative assessment are 'not necessarily the highest grades available or even directly related to competent practice [...] provided the work is (eventually) completed, the award will be achieved' (Torrance, 2007: 284). More dangerously in terms of assessment and its effect on the teaching and learning process, the findings of a study by McNeil & Valenzuela (1998) conclude that 'behind the rhetoric of rising test scores are a growing set

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

of classroom practices in which test-prep activities are usurping a substantive curriculum'. (McNeil & Valenzuela, 1998, p. 2). Torrance (2007) calls this the 'displacement' of learning (i.e. understanding) by procedural compliance: i.e. achievement without understanding' It refers to assessment that masquerades as, or substitutes for, learning itself' (Torrance, 2007: 293). This 'displacement of learning' is now a firmly ingrained part of modern assessment practice worldwide and this displacement, despite the apparent goodwill of the other test stakeholders (the teachers and parents) in helping students to pass, cheats everyone involved in the process.

What is clear from the research mentioned above is that the often high-stakes nature of summative assessment and testing is influencing the teaching and learning process. Teachers are either consciously or unconsciously directing their teaching to help students pass their tests and get good grades at the expense of other areas of learning, and the students are focusing too closely on the processes of passing these tests, perhaps at the expense of their own general personal development or going against their natural learning style.

When considering the negative influence summative assessment may have on the teaching and learning process, researchers suggest that for EFL contexts at least, there may be large gaps between the 'communicative competence' of the speaker, and the kind of competence needed to gain high scores on the test, a kind of competence usually termed 'strategic competence' – a term Canale and Swain (1980) define as 'verbal or non-verbal communication strategies that may be called in to action to compensate for breakdowns in communication due to performance variables or to insufficient competence' (Canale and

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

Swain, 1980:30). Bachman (1990) claims that knowledge of strategic forms of language competence will influence test performance. Bachman uses the theoretical example of two students taking three tests, with the first test being a test of language usage, the second being a test of receptive performance, and the final being a productive oral performance test. The marks from the first two tests may be identical, but one student may get a lower score than the other on the final test because that student was not so aware of the strategic techniques used to get higher scores in the event of a gap in communicative competence. Bachman suggests this kind of event would show that while the two speakers language competence may not differ so much, it was the knowledge of and use of strategic competence that caused the higher-performing student to ‘differ in their willingness to exploit what they knew and their flexibility in doing so’ (Bachman, 1990:105), and so receive a higher score on their test. If we imply for a moment that teachers are aware of the necessity of strategic competence and either consciously or unconsciously attempt to train their students in the forms of strategic competence necessary to get high marks in a test, then the test could be said to not give a fair representation of their students’ true communicative competence, and the validity of their grades would be called into question..

With all this in mind, summative assessment (that of assessment ‘of’ learning as outlined above) has raised a variety of complex issues within different contexts. However, similar issues also arise when we consider the effects that formative assessment – that of assessment ‘for’ learning – has on the teaching and learning process, particularly when the formative assessment takes on a criterion-based approach.

Dangers of Criterion-Based Formative Assessment

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

If assessment 'of' learning can be said to be problematic, what then of assessment 'for' learning? As outlined above, formative assessment is supposed to provide a beneficial source of assessment that provides feedback to teachers and learners about their learners' progress, spotting areas of trouble and guiding future classes or revision of the curriculum. However, if the teacher or educational body in question is guided by a pre-defined criterion approach to assessment 'of' learning, then any teaching preparation or feedback from assessment 'for' learning that takes places before, during and the after the assessment period is likely to be influenced heavily by such criterion at the same time. Torrance (2007) explains the situation as follows:

'This might be characterized as a move from assessment *of* learning, through the currently popular idea of assessment *for* learning, to *assessment as learning*, where assessment procedures and practices come completely to dominate the learning experience, and 'criteria compliance' comes to replace 'learning' (Torrance 2007:282).

The idea of assessment 'as' learning is finely engrained into the curriculum of many EFL programs and textbooks, with most teachers' guides in EFL textbook programs containing 'learning objectives' for the week, month, year, semester, unit, module or other period of study. Teachers use these learning objectives to shape their lesson plans, prepare materials, and shape to an extent the communication from teacher to student in class, as the teacher tries their best to make sure the students meet their learning objectives for that day. However, the effort required to meet these learning objectives may often come at the expense of other student needs, whether they are communicative, strategic, linguistic or otherwise. Torrance continues

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

What the learner can do, and can be seen to do, in relation to the tasks required of them for competent practice, are paramount. It is of little interest to the learner or assessor to identify what *else* the learner can do (i.e. engage in divergent assessment) although this may be of considerable importance to their longer-term development. (Torrance, 2007:292)

In an EFL context, if any formative assessment done by teachers is affected by their meticulous attention to these fine-grain learning outcomes that supposedly constitute ‘learning’, then that assessment may be missing the point of what it was intended to bring on a general level - that of increasing the overall linguistic competence of the students joining an EFL program. Torrance mentions that this practice is commonly followed in many assessment contexts, due to ‘the long-term advocacy and development of criterion referencing and competence-based assessment in the post-compulsory sector, and partly in the attainment-oriented culture of the sector.’ (Torrance, 2007:292).

In EFL contexts, criteria-based assessment is further problematic in terms of the reliability of formative (as well as summative) testing. Usually for formative assessment, testing comes in the form of tasks that are focused on the area of linguistic competence to be assessed. What is difficult here is that when teachers make tests for their students, they have to be sure ‘that the given sample of items or tasks is relevant to that domain’, or the domain score (whether the student has fulfilled the criteria or not) that the student was supposed to get under that domain will be incorrect. (Bachman, 1990:212). When teachers design tests, they may often be confused about what the criteria for assessment may be, in terms of the wording of the criteria in the curriculum, or whether any tasks the

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

teacher makes from those criteria are actually relevant in practice to that domain at all, either before or during a task, at the feedback stage, or in post-task follow up activities.

Another problem occurs when defining a student's mastery of a particular section of the criteria on the curriculum. For an EFL context, this would mean trying to ascertain the exact mastery of the target language form or area of communicative competence in the terms laid out in whatever EFL curriculum the EFL program in question decides to use. To do this, teachers usually put in a 'cut-off score' (Bachman, 1990, 215) during testing that shows whether the students have 'mastered' the target forms or not (usually given as a percentage). However, the difficulty then comes with distinguishing the abilities of students who have met the minimum requirements for 'mastery' of the target form by meeting the cut-off score, but have different percentages over the cut-off score nonetheless. For example, if one student has a test score of 100%, and the other has a score of 70%, but is still above the a supposed cut-off score of 60%, then it may be the case that while the two students may be said to have mastered the target language form in question, one has 'mastered' it a little more than the other. However, under formative assessment – the kind of assessment that should dictate follow up activities based on the students' scores, successes and failures – these two students may be treated as being at the same level, with the slightly underperforming student receiving no advice on how to boost his score to 100% as he or she has – according to the teachers' adoption of criteria-based assessment – already 'mastered' the target form in question. The case could be made that the exact same situation could also occur in our context, as we too had introduced a cut-off score of 60% from the three-month-average test scores as a marker of success or failure. It would potentially be problematic to assess the success of one student who 'passed' at 60% and

another student who 'failed' at 59%.

After reviewing the above literature concerning the dangers inherent in certain approaches to assessment, I realized that while the new EFL curriculum I had created at the foreign language education centre was certainly an improvement over the last year's ad-hoc system, there may have been some theoretical and practical issues with its implementation. However well-intended and apparently, on the surface, good practice this new criterion-based curriculum and its subsequent assessment procedures seemed at the time, it was possible that the introduction of this kind of curriculum into a children's EFL language program may have presented the program with unforeseen consequences of their own, and I decided that these consequences would warrant further investigation.

Research Rationale - Method

To investigate the impact of assessment on this EFL program, I decided to conduct a case study of the students, parents, and teachers who were involved, in order to get an insight into the impact from all the stakeholders in the program. I developed the rationale for the case study from two published studies into the impact of assessment on the different stakeholders involved, namely Deforges, Holden and Hughes (1994) and Scott (2007). The first study (Deforges, Holden and Hughes (1994)) focused on assessment at Key Stage one level in schools in the UK. They found that:

- '(i) there was substantial variation in parents' views concerning formal assessment; (ii) teachers were mostly unaware of parents' views on assessment; (iii) the assessment and reporting processes had relatively little effect on parents' views or on teachers'

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

perceptions of parents' views' (Deforges, Holden and Hughes, 1994:132)

This research set out to answer questions regarding assessment with reference to the perceptions and actions of parents and teachers, namely their initial views, how these views influenced their behavior, and how the interaction between parents and teachers affected the learning process, all factors that I could replicate in this research.

Scott's (2007) study was similar to that of Deforges, Holden and Hughes (1994) in that Scott's study on stakeholders' perceptions of test impact used interviews of the stakeholders to explore their perceptions of statutory tests and their potential impact on the learners. Her findings suggested that 'there is a need for caution in stating that the tests have value for diagnostic purposes, other than over a cohort, or in reporting to parents' (Scott, 2007:44). Scott used an inductive analysis of interview data with learners, teachers and parents in the UK, focusing on learners who were not native English speakers and who were learning English as an additional language (EAL). Scott did this by searching the transcripts for insights into participants' perceptions of and attitudes towards assessment in relation to their impact, as well as references to washback on the teaching and learning processes. It was hoped that by emulating this study I would gain useful information on the stakeholders' perceptions in my context.

It was hoped that by conducting my interviews with a similar approach to the ones in Deforges *et al* (1994) and Scott (2007) that I could find examples of impact on the stakeholders from the negative aspects of assessment in the EFL program's new curriculum as laid out in the literature review of this research. As with Scott's and Deforges, Holden and Hughes' research, data collection was done by way of one-on-one interviews with

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

parents, teachers and learners at three stages of the assessment process. Firstly, I held interviews at the beginning of the school year, at a time after the teachers, learners and parents had been informed of the new assessments and had been given a copy of the first months' provisions of the criterion-based curriculum, but before the first round of assessments had actually taken place. I held a second round of interviews at the end of the second month, by which time the students had completed two rounds of assessments and were settled into the new system, the teachers had gotten used to the preparation, provision and aftermath of the assessments, and the parents had received two sets of monthly reports with accompanying test grades. The final round of interviews were held after the third month of assessment, by which time the students had received their average grades over the three-month period, as well as had received the class average grades over the three months for comparison of their performances against the other students in the class. Any students who got test scores of under 60% at this time would have to be re-level tested and probably moved to a different class, so this was a critical time to gauge underperforming students reactions, as well those of those students who were under pressure from their parents to succeed under the new assessments.

Participants:

The case studies centered on four learners, with the English nicknames Claire, Suzie, Jane and Ben. Claire and Jane were in the same class, as were Suzie and Ben, with both classes being in the third highest level of the program. Each student was in grade 5 of elementary school, meaning they were all aged 12 years old at the time of the interviews. They were selected for this case study as I felt they were old enough and aware enough to give insightful responses to the questions posed at interview, having taught each of them at Crosthwaite, P. (2009). *The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. Asian EFL Journal. M.A. Thesis.*

various times over the last year. These students had also been coming to the language centre for over a year and were therefore familiar with the old ad-hoc system, meaning that they could comment more meaningfully on the changes that were to be implemented.

The parents were all available to come into the language center to pick up their children and thus were usually available for interview. They were keen to join in the interviews as it was the first time they had ever been consulted on their children's academic progress in this way by the language centre, and thus they felt like they were taking a larger interest in their children's progress.

The teachers involved in the case study were the teachers who were directly involved in teaching the four students in question, with teachers Damien and Michelle involved in Claire and Jane's class, and teachers Emily and Keri involved in Suzie and Ben's class. Damien and Keri are native English speaking teachers from the U.S.A, and Michelle and Emily are native Korean English teachers. The Korean teachers and the native English speaking teachers each taught the same class on alternate days. The teachers had all worked at the language centre under the old conditions for over a year.

Approach

For the interviews, I was present at each stage of the interview process. Interviews were conducted in English with the native English-speaking teachers and in Korean with the native Korean teachers, students and parents. I generated the first round of interview questions to be held in March 2008 by reviewing the research literature (namely Deforges, Holden and Hughes (1994) and Scott's (2007) research, as well as Sadler (2007), Torrance Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

(2007) and Harlen and Crick (2003)) and formulating questions that sought to explore whether their findings would also be found in my context.. The second and final rounds of questions sought to probe deeper into the interviewees' initial responses and to see whether and to what extent these initial responses had changed over the three interview periods. All the questions asked in the case study can be found in appendix 3

The questions for the interviews were translated into Korean for the purposes of the learners and parents by two assistants at the language institute, who were both briefed by myself on the investigative purposes of the questions to be asked and could therefore translate the questions more accurately. This may have been one of the limitations of the interviews, in that I could not be 100% sure that my questions (or subsequent translations of the Korean transcripts) would be accurately translated, or that the meaning or tone of the questions or translations could be misconstrued, despite both my assistants' high English language proficiency. The interviews were held with myself, the interviewees, and my supervisor who would ask the questions to the learners and parents in Korean. The interviews were held in an empty classroom, with the parents and learners waiting outside in another classroom until they were each called in for interview. My supervisor interviewed each learner and parent at each stage of the process. The interviews were recorded on tape and the answers were transcribed verbatim and translated into English firstly by my supervisor and then by my Korean wife, who met to compare their translations in order to decide on the most accurate version of the transcriptions. Another limitation was that as I could only review the interviews after they were recorded and transcribed, I had no way of asking immediate follow-up questions that would probe deeper into the interviewees responses, meaning that I would only receive answers to the questions

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

that were on the interview questions sheet I had given my supervisor. I feel that this made the interviews more shallow or superficial than they would have been if I could have conducted the interviews myself in English, especially during the learner's interviews as they often gave one word answers such as 'yes' or 'no' and I wasn't able to get them to explain the reasoning behind their comments.

In order to prevent influencing the answers of the learners, the stakeholders were only informed at the end of the final interview process about the true nature of the research, and all the interviewees gave their permission for the interviews to be used. Some small mention of the purpose of the research was given to my supervisor in charge of asking the questions in Korean to the learners and their parents, yet I didn't fully divulge to her my expectations for the research in a bid to avoid influencing her in a way that might damage the results. However, in hindsight, there was not enough time to consider the validation of the interview procedure as outlined in Alderson and Banerjee (2001), in that the order of the questions may have limited the interviewees' responses, and that the interview questions did not mean the same to the interviewees or to the interviewer as they did to the researcher.

Once all the interviews had been completed, I looked at the transcripts of the recordings and attempted to see if there was any correlation between the interviewees' answers and the findings of the research that I put forward in the literature review section of this paper. In my commentary of the results, I was careful to also include sections where there was little or no correlation between the interviewees' responses and the research's findings, in order to highlight any differences that may have been found in the unique

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

context in which these interviews were held.

RESULTS

In order to show how the perceptions of the stakeholders towards the new system of language assessment developed over time, I have divided the results of the case studies into three sections, relating to the three periods of time that the interviews were held. The first section discusses the interviews held before the introduction of the new assessment and curriculum in March 2008, after parents first received copies of the new syllabus, and were informed as to the implementation of the new assessment system. The second section discuss the period of the second interviews, held after the second round of assessments in April 2008. The final section discusses the interviews held after the third month of assessment and after the average grades and reports had been released to parents at the end of May 2008. A complete transcript for Claire's mother's May interview can be found in Appendix 4 as an example of how the interviews usually progressed.

First Interviews – Pre Assessment – The Learners

The initial round of interviews for the learners focused on the perceived importance of the assessments as well as to find out if there was any pressure already on the students from the new monthly assessments that were to take place. Furthermore, the interviews

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

attempted to find out if the students had any knowledge about how the assessments would be conducted and whether they had any understanding about the meaning of the grades they would be given.

When asked about whether or not there was any undue uneasiness or stress caused by the impending assessments, the students gave mixed answers. At the initial stage of assessment at least, the assessments had given some of the students cause for concern, even though the assessments are only part of their EFL program and are outside the standardized tests of their public school system. According to Scott (2007), There is ‘concern regarding the potential impact of tests on the children taking them in terms of inducing undesirable levels of anxiety and the potential for long-term effects on self-esteem, motivation and later academic success or failure (Kirkland, 1971; Clarke, 1997; Pollard *et al.*, 2000, in Scott, 2007:27-28). It would remain to be seen if Jane and Ben’s responses might change over time.

Next, the students were asked about whether pressure over the assessments may come from their parents, in line with the belief that assessments ‘increase the pressure on students to do well resulting from the aspirations of parents and teachers (Davies & Brember, 1998; Leonard & Davey, 2001). Only Suzie mentioned that ‘I get a lot from my mum’, with the other students claiming their parents hadn’t pressured them over the assessments at this initial stage. This was quite surprising to me given my experience with Korean parents and their children in terms of their competitiveness in the past, and I would have to wait and see whether the pressure increased as the assessments continued.

The learners were then asked about their perceptions of the percentages that would be given out after every assessment to get a feeling of their views on summative assessment. Each student believed that the test scores would be a good measure of their foreign language ability. When asked about the importance of the tests and the possible validity of the new assessment system for checking their achievements, the students were a little vague, with only Claire mentioning that ‘the tests are important to check a learner’s ability’, and the other students saying the tests were ‘just important’, with Jane mentioning that she didn’t know well whether they were important or not. At the pre-assessment stage, it is therefore possible that students were not aware of the testing procedure and the grading process. When pressed on the purposes of the assessments, the students eventually replied that they didn’t know well about the new assessments and why they were being implemented. Evans and Engelberg (1988) used a questionnaire to study students’ attitudes to, and understanding of, teachers’ grades. In terms of the understanding of grades, they found that ‘older students understood simple grades more than younger ones, but even older students did not understand complex systems of grades.’ (Evans and Engelberg (1998), in Harlen and Crick, 2003:187).

Even at this early stage then, there appeared to be some uncertainty into how much of a formative benefit the assessments would bring, as well as a level of anxiety regarding the assessment process for these learners. We would see later whether this uncertainty or anxiety would continue in later months.

First Interviews – Pre Assessment – The Parents

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

As with the learners, I was keen to get the parents' opinions on what the new assessments would mean in terms of the grades they were to be given. The parents, like the learners, believed initially that summative assessments would give a clear indication of their child's progress

(CLAIRE'S MOTHER) 'It's the result of what we have learned'

(SUZIE'S MOTHER) 'It's an evaluation of the in-class behavior of the learner'

(JANE'S MOTHER) 'It's the result of the learners understanding'

(BEN'S MOTHER) 'It shows the correct judgment of the student'

As with the learners, however, when it came to how the assessments were performed and graded, as well as understanding the information gained from the assessments, the parents were unsure. When asked about how the assessments would be conducted, Suzie's mother, Jane's mother and Ben's mother were all unsure about the details of the assessments, even though they had been informed by correspondence by the administrative department a couple of weeks prior to the first set of interviews. The parents (with the exception of Claire's mother) were also unsure about the differences between the six levels in place at the EFL program (separated by performance on our new criterion based curriculum).

When asked what kind of information the parents hoped to get from the assessments, most parents hoped to learn about their child's understanding of what was taught in class. The parents believed that even though they were as yet unaware of what may come out of the assessments, they believed the new criterion-based approach to be a

good thing, with Ben's mother going so far as to say that 'this is the best system they have introduced, I believe'. Furthermore, every parent liked the implementation of the criterion based system. They all felt the assessments to be necessary, in keeping with the finding of Scott's (2007) case studies, where she mentions:

Parents potentially accommodate contradictory notions of statutory assessments: (i) that they are evaluative, transmitted through references to 'average' levels, and (ii) that they are diagnostic, transmitted through references to what children need to do (Scott, 2007:38)

The parents also preferred to receive teacher's comments about their student's progress over a report containing test scores (with the exception of Suzie's mother), stating a dislike for summative assessment reporting and a wish for a more formative approach to assessment and reports, with Jane's mother stating 'I trust the teacher's objective comments more because their test results are quite changeable depending on the students' condition on the test day.' This view is also supported by Shepard & Bliem (1993), who found that, although parents were interested in test results, they consistently reported that teachers' reports of their children's progress were more meaningful and that they had greater faith in those. Yet Scott's (2007) research would seem to disagree with the parents' enthusiasm for formative assessment.

'Although both test results and teacher assessment are reported to parents, since only test scores are reported in performance tables, there would appear to be merit [...] that test scores bear more weight than teacher assessment in the public arena' (Scott,

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

2007:37).

Next, we asked the parents about the possible anxiety, stress, or motivation the new assessments might bring. Three of the parents firmly believed that the assessments would stimulate their children to learn and perform. However, this new found motivation did not come without some pressure. When asked if they thought the learners would feel any pressure from these assessments, the parents gave mixed results:

(CLAIRE'S MOTHER) 'I think they already feel some pressure'

(SUZIE'S MOTHER) 'Yes' (the students will feel pressure)

(JANE'S MOTHER) 'It depends on the student but my kids don't feel much pressure'

(BEN'S MOTHER) 'When they study well and prepare well they won't feel much, but if they don't study well they will feel pressured'.

There didn't seem to be much evidence at this time that the parents would pressure their children in any way to perform well under the new assessments, yet the parents perceptions of what may happen to their children if they underperformed were interesting, with Jane's mother claiming 'when they get a good score, they get a feeling of achievement and want to maintain that good score but when they get a bad score, they get very disappointed. This is a worse feeling than getting a good achievement'

The final thing I was interested in knowing were if these new assessments would promote teamwork between parents and teachers and the language institute. When asked about the possibility of improvements in the parent/teacher working relationship, the parents were all hopeful as to a positive outcome. As for the parent's perceptions of the role
Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

of the teacher, the parents believed that it was the teacher's responsibility to lead the learners, with the parents existing in a supporting role.

First Interviews – Pre Assessment – The Teachers

The Korean English teachers (KETs) were positive about the introduction of the new assessments, but were mixed about the formative and summative aspects of them. Michelle stated 'We need to suggest a new system and curriculum not annually but during a shorter period. The students settle down in their work and can also change their attitude in class not only to parents but also teachers', whereas teacher Emily saw a more summative approach, stating 'we can know how well they followed the curriculum's guidelines when we get the test scores'. However Emily is also aware of the formative aspects of the assessments, stating 'If there are students who can't reach these goals, we can see clearly which points they are not following well, so it is helpful for teachers and parents. For the native English speaking teachers (NETs), Keri and Damien were also positive, but again with Keri taking the view that the assessments would be useful in a summative way, when she said 'hopefully, the new system was introduced to standardize the grading method among the teachers as it was pretty ambiguous at times, and with the speaking tests, ideally we are all grading from the same page using the assessment criteria.' Damien focused on the formative aspect, when he mentioned 'I feel that the grades will represent how I proceed with the rest of the lessons, and act as a gauge for me to see how strong they are, and it will allow me to modify my approach in further lessons.'

Both KETs made it clear that it was also their responsibility to make the sure the

students were kept informed as to the meaning of the assessments. Scott (2007) mentions that ‘there is a tension here between ‘protecting’ children by keeping them unaware of the importance of the tests and informing them in order to try to elicit best performance.’ (Scott, 2007:41), but at this early stage it would appear the KETs have chosen the latter. For the NET’s, Damien agrees in part with the KET’s comments, wanting to ‘have a class to explain it to them’, in keeping with the KET’s approach. Keri, however, was more negative, claiming that the students and parents had no idea what the grades meant, saying ‘The focus here is on the numbers rather than understanding what the scores mean, for example, what is a 10, what is a 1 etc. I even wonder how much the parents are aware, outside of seeing a little number circled on a report.’

In terms of anxiety or stress caused by the assessments, Michelle, contradicting herself a little, is mindful of misjudging the results of the tests, stating ‘that’s why it can be bad to judge students by tests because we can’t consider emotional factors during the test’. Michelle, despite appearing on the surface to not hold much stock in the test results due to emotional factors, is also very mindful of the stakes of the tests for her as a teacher at the language institute. Damien hinted that the anxiety that may be caused by the tests may actually be a good thing, stating ‘I think it will keep the kids on the ball and they will be even more aware that there is a test coming up, which may push them to try harder.’

For both KETs, closely following the criterion based guidelines in the assessment is paramount, with Michelle going so far as to say ‘this system will be really effective if teachers understand well and train students enough they can study well after that. Both teachers then could be said to have already consciously or unconsciously shaped their

classroom activities towards following the criterion-based assessment procedure. The idea of training here mentioned by Michelle is very important. This is echoed again when Michelle, answering a question about the meaning of the tests, stated ‘We need to grade the student’s level from the result of student’s study in class. Then we can train students up to standard for the next test.’ Sadler (2007) mentions that this practice has ‘become so elaborate, and the level of assistance so comprehensive, that the learner cannot help but ‘succeed’ (Sadler, 2007:390), an opinion supported by Torrance (2007), who states ‘such involvement helps to develop teachers’ understanding of the assessment process and criteria which, in turn, they pass on to students through exam coaching.’ (Torrance, 2007:285). However, it seems the teachers were attempting to find the tradeoff between following a strictly summative approach and finding a useful formative outcome to the assessments at this point. Whether or not Torrance’s suggestion that actual ‘coaching’ might take place would remain to be seen at this point.

The KETs believed that the test scores were useful in assessing learning. Keri was much less optimistic than the KETs, stating ‘I think it can be ambiguous at times, as if little Johnny can understand what is in the book, but can’t express his thoughts well in English conversation, then he will be graded lower, but little Annie, who can’t speak English well but can regurgitate the right answers will be graded higher.’ This was at odds with Damien, who said ‘As long as the tests contain material that has been done by the students then I think they are useful’. I believe the KETs here were trying to speak of the new system in a positive light, for reasons to be outlined in the final interview section of this paper. Keri was much less optimistic than the KETs, criticizing both the criterion-based curriculum as well as the importance of the speaking tests at the expense of other

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

areas of English competence such as reading, writing and listening. This was at odds with Damien, who was more upbeat in his appraisal of the meaningfulness of the assessment grades, when he said ‘as long as the tests contain material that has been done by the students then I think they are useful.’

Finally, although the KETs both agreed the parents would be interested in the new assessments, they were mixed about the parents’ role for their children through these assessments. However, the teachers’ perceptions of the parents’ interest in the assessments go against Deforges, Holden and Hughes’ (1994) findings that ‘several of the teachers made explicit reference to what they saw as parents' lack of interest in the subject: 'They don't seem that interested. They just accept it as something that is happening.’(Deforges, Holden and Hughes, 1994:142) This follows Keri’s somewhat pessimistic view of the interest of the parents about the new assessments. She explained:

‘I think some are and some aren’t. In the past, when we had problems with the kids, we would call and the parents would just kind of wash their hands of it, so I’m not sure the parents will understand the new curriculum if their kids aren’t speaking English at home, which they probably aren’t.’

Clearly, the KETs and the NETs had different views regarding the effectiveness of the assessment process, the benefits of the new criterion based curriculum, and the role of the parents in terms of supporting the learners. Both KETs were quite optimistic for the new assessments, but as mentioned before, I felt Michelle and Emily to be giving overly positive answers for reasons to be outlined in the final interview section of this paper.

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

Damien held similar upbeat views, yet the differences between his positive outlook and Keri's negative outlook are quite striking. This could be explained by Keri's longer teaching experience at this particular language institute, therefore perhaps she felt more comfortable than Damien in criticizing the program.

Second Interviews – Mid-Assessment period – The Learners

The second round of interviews were recorded at the end of April 2008, shortly after the learners had taken the second round of tests and after the first set of results had been made public to the parents. These interviews focused again on test-related anxiety, as well as the impact of the learning objectives from the criterion-based curriculum, and finally on whether the teachers or parents had tried to influence the learners' performance in any way.

In terms of anxiety and motivation, the learners were still divided on whether the assessments created anxiety, with Claire and Suzie again expressing that they felt anxiety, and Jane and Ben again mentioning they didn't feel any undue stress. However, when asked again about her levels of motivation for the assessments, Jane mentioned that 'if I don't get a good score compared to the other students then it's not good, and if the score is not good my mum will kill me'. When asked what the consequences may be if learners fail to meet the provisions in the learning objectives, Suzie and Jane mention that they will

be punished by their mothers, with Suzie suggesting that the teachers may punish her also, acknowledging that the other stakeholders in the assessments may cause her discomfort if she fails to perform at the level expected.

The students all believed the learning objectives to be useful in some way, as were the tests, but only gave 'yes' answers when asked about the positive impact of both. The students felt that the results of the tests highlighted their weak points, but their answers were unrelated to the learning objectives, with the students mentioning that memorizing vocabulary was their greatest weakness. It seems then the students were not able to understand or had not read the criteria themselves, and were oblivious as to the criteria's content, instead focusing on the usual day to day problems of rote-learning vocabulary.

Finally, the learners didn't see much of a formative aspect to the assessments after the test period. When asked if their scores were used to guide the learning process in the post assessment phase, they answered negatively. This lies in contrast to the KET's as well as Damien's earlier assumptions that the tests may have a formative purpose, as we can see by the learners' answers that so far the tests have not had any formative purpose (for the learners at least) at all.

Second Interviews – Mid-Assessment period – The Parents

The parents at this stage all mentioned that they were pleased with the results their children had received so far. When asked again, as with the first round of interviews, whether the test results had any meaning to them personally, they trusted that the results of

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

the tests had been accurate in measuring their child's abilities. The parents also decided that the learning objectives were useful in assessing what had been learned, and that they were useful in aiding the learning process.

The parents believed that when a child had passed the weekly learning objective, the student had gone on to 'master' the provisions in that objective, as long as the teacher 'led' them well enough. This is in contrast with Bachman's (1990) views on 'mastery' of the criteria outlined in the literature review in terms of the 'cut-off score' given by teachers to get a passing grade, where two students who pass may have varying levels of mastery of the target form. They were, however, less optimistic about what would happen to those students who did not perform well in class and could not 'pass' the weekly objective.

In addition, Claire's mother, Suzie's mother and Jane's mother expressed doubt that the speaking tests had any formative purpose, in disagreement with their earlier comments. When questioned as to whether the test results were used to guide the learning process, they stated:

(CLAIRE'S MOTHER) No

(SUZIE'S MOTHER) Yes, we get the results, so I know their level. Only using the scores to stimulate the students is not a good idea but there is nothing we can do. We just need to motivate them to get them better I suppose.

(JANE'S MOTHER) Yes I know, I don't base it all on that test

Claire's mother, in a later question, went so far as to say 'we can't just simply use

the tests to show well what she has learned.’

In terms of how the parents and teachers tried to influence the students for the purposes of the assessments, they expressed that they needed to do more, despite their children’s claim that their parents had tried to help them with their assessments. As for the teachers’ influence, the parents claimed that the teachers had heavily influenced their children during the assessment period. Perhaps in this case then, the children had expressed to their parents some of the pressure that the teachers may have put them under to perform well. The parents, with the exception of Ben’s mother, acknowledged that their children were feeling undue stress or anxiety from the assessments.

The parents all claimed that they were working together with the teachers to aid their students’ performance in the tests, and that they were more aware now of the importance of the tests and what they involved.

Second Interviews – Mid-Assessment period – The Teachers

The teachers were asked very similar questions to those of the parents at this stage of the interviews, in order to see if there were any discrepancies between the perceptions of the teachers and parents over the value and impact of the assessments, and of their roles as stakeholders in the assessment process.

The overwhelmingly positive response to the introduction of the new assessment system had faltered somewhat for the KET’s. Emily felt that the tests did not cover

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

everything that had been learned in the classroom, saying ‘sometimes, some students are good at writing or vocabulary, however unexpectedly they are not good at speaking to the English teacher or might be nervous or misunderstand the question so their scores weren’t that good enough’. Both teachers said that the tests were not fulfilling their supposed summative roles in judging any ‘assessment of learning’, with Michelle mentioning that ‘at the moment now we can’t judge the meaning of the results only with that score’, and Emily mentioning that ‘I don’t think the tests involve all the things that students have studied, so although they are graded correctly, it is difficult to check their true level’. For the NET’s, Damien was still upbeat about the grading and felt it was useful for him, but Keri took a similar view to Emily in suggesting that in terms of ‘assessment of learning’, the tests had not been useful.

This dissatisfaction spread also to the learning objectives, with Michelle saying ‘in the case of the learning outcomes, it is difficult to judge every ability’, and Emily saying ‘the rest of the stuff all depends on the teacher.’ The KETs were now saying that the learning objectives were more useful for guiding what should be taught rather than assessing what has been learned in the classroom. The NETs complained that the learning objectives were in fact too difficult for many of their students, especially those who had difficulty with writing, as many of the objectives in the curriculum expressed that some degree of competence in writing was necessary to achieve a passing grade, yet many students were unable to write sufficiently well. However Damien still remained positive, mentioning that the objectives gave him lots of ‘conversation potential’.

In terms of undue stress or anxiety from the learners regarding the assessments, the

KETs gave mixed answers, with Michelle believing that stress was caused, and Emily disagreeing. Both KETs were positive in terms of the students' motivation under the new system, but this seemed to stem more from the colorful new textbook rather than by the assessment process itself. The NETs were mixed in the seriousness of the stress in their response, with Keri mentioning that she 'they stress out when they know there is a speaking test, but we started doing spelling tests in my class too, and they usually cry if they don't do well', but with Damien mentioning only that 'kids always get disappointed when they have tests'.

Probably the most important findings from this round of interviews were that there was evidence from the KETs that echoed Sadler's (2007) findings that teachers were guilty 'of getting students through – often at the expense of what it really means to learn' (Sadler, 2007:387) and Torrance's notion of 'displacement of learning' (Torrance, 2007:293).. Both KETs, when asked about what would happen if a student could not meet the goals of the learning objectives, mentioned that they force the students through a variety of means to 'pass'. Emily mentions:

'Actually we force them to fit in, or lead them. To be honest, if the student isn't at the right level, we can just give them more homework or give them one more test, so we try to fit them in.

Damien also offers more learning activities, or simplifies existing ones, if the learning objectives are not being met in that week, saying:

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

‘Often we find alternative activities, like I’ve had some students become a little bit emotional about not being able to do something, and so I’ll often scale back the activity and make something more manageable for that particular student.’

Keri is much blunter than Damien in this issue however, saying:

‘I think unfortunately they get pushed through. Like, I have kids in my K-6 and IS3, IS4 who can’t read. On the speaking test, I will say, like one of my questions will always be ‘read this’, and the students can’t read, and I don’t think there is anything done to encourage that, like a remedial program.’

In this case then, the assessments could have said to have totally failed to serve any kind of formative purpose as the learning objectives are simply being disregarded if the students cannot meet them.

Michelle and Emily give a mixed response regarding the parents’ influence on the performance of the tests, with Michelle seeming to have more contact with parents, and thus expressing a more positive view, but with Emily unsure about what happened after the test results got home to the parents. It did seem like teamwork was taking place, yet this was strictly for the KET’s only, with the NETs almost totally shut out of the parent / teacher discussions, presumably due to the language barrier. However, this did not stop Keri giving her opinion on how parents influence their children during assessment, and she mentioned:

‘I think they try to, but I don’t think it’s a positive influence. I think its pressure which makes them cry, makes them crack. In Korea, maybe I’m wrong, but my impression is it’s not about making sure your child is learning, its about what level is your child, like ‘as long as my child is progressing’, they don’t really care if they are learning.’

Certainly here, and in my experience also, there is a serious discrepancy between what NETs think about the influence of Korean parents on their children’s learning, and what the KETs or parents themselves think about their influence. NETs in Korea generally believe the competitive nature of Korean parents means that they are only concerned with their children getting good scores, at the expense of all other emotional considerations, an opinion definitely held by Keri here, but not to much of an extent in Damien, whom it must be noted, is the less experienced teacher of the group.

Final Interviews – The Learners.

The final interviews took place in May 2008 at a time when the students had completed their third round of assessments, and the results of which were sent home to the parents.

In terms of anxiety or pressure from the assessments over the three month period, the students all believed that the pressure on them to perform was mounting, and that the source of most of that pressure was their parents. This was in contrast to Jane and Ben’s statements at the beginning that they did not feel any pressure. The students mentioned:

(CLAIRE) When I don’t want to study for the tests, my mum and dad try to force me to do
Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

it. I feel a lot of stress.

(SUZIE) In the beginning, my parents wanted me to study a lot for these tests, but because my vocabulary isn't so good, they haven't pushed me that much recently, so I do it by myself.

(JANE) She [my mother] doesn't let me sleep at night so I always wake up late. It's very stressful.

(BEN) After the tests I don't want to hear her [his mother's] voice ...it's too much.

When asked whether the students felt any pressure from their teachers, only Jane claimed she felt pressure, which would suggest again that the parents seem to be the main source of stress regarding the assessments rather than the teachers.

Regarding the perceived value of the assessments over the period of the case studies, the students still felt that the tests were useful, but perhaps not in the way I had intended when I designed the program of study and selected the assessment approach. Three of the students suggested that the main benefit of the tests was that they concentrated more on vocabulary, so their ability to remember lots of vocabulary improved. The students were also split when asked if the new system had motivated them to study English further. It seems the tests were only forcing students to memorize vocabulary, rather than to be a useful diagnostic tool in determining whether or not the students had mastered the provisions of the curriculum, which could be said to be evidence of Torrance's (2007) 'displacement of learning', or 'achievement without learning' (Torrance, 2007:293).. With this in mind, the tests could only be said to be of little value and in my opinion were becoming counter-productive in that they were only making the students focus on

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

memorizing vocabulary.

In keeping with the students' interviews in April, the students again downplayed any formative benefit from the assessments. When asked about the kind of feedback they received after the assessments, Suzie and Ben enjoyed some reward for getting a good score, but Claire and Ben mentioned their feedback was quite negative:

(CLAIRE) Our teachers told us to do better next time, and my Mum, she said 'how can you get a score like this?', and my dad said 'this time the test score is poor, so next time don't perform like this.'

(BEN) I didn't get any good feedback. My mum doesn't give me much feedback but I can feel she is very cold to me about 2 days after the test

According to the students, any feedback that did occur tended to be linked with success or failure on the tests. Success brought with it positive comments from teachers and parents (as well as the possibility of some reward in the form of a present or snack), and failure brought with it some form of punishment. The students did not mention any kind of feedback received that may have served a formative purpose, such as how to correct any mistakes they had made or guide them in what to study next. Three of the students did believe that the reports written by their teachers after the assessments were useful in some way (giving a 'yes' answer when asked if the reports were useful), but were unclear as to exactly what kind of benefit they provided, with the exception of Jane who valued the reports a lot as her mother would buy her something nice if her report was positive. Thus far then, in the students' opinion, neither the tests, curriculum nor reports at this stage

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

provided any formative benefit.

Finally, again in keeping with the findings of the second interviews, the students continued to suggest that they were being conditioned to take the tests by their parents and teachers, with Ben going so far as to suggest that ‘our teachers trained us by showing us what the important things are that may be on the test [...] if we remembered those words, we could pass’ and Suzie mentioning that ‘my parents and the teacher affected us all a lot. They created the environment for study, and we just followed them.’ One can see here the power of the other stakeholders in the assessment process and the effect that they have on the learner, particularly if the learner is young and therefore more easily influenced by teachers and parents. If the teachers are training their students for the tests in the way Ben describes to help their students perform better on the tests, then the tests could not be said to a true representation of the students’ mastery of the objectives outlined in the curriculum.

In contrast with the students’ earlier belief that the tests allowed to them to correctly gauge their ability, when the students this time were asked whether they believed the tests were a fair representation of their true ability, they were not so positive.

Overall, the students appeared far less positive about the benefits of the assessments and the curriculum than they did initially, and the influence of the other stakeholders appeared to be influencing their behavior more strongly than before.

Final Interviews – The Parents.

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

The parents were still positive after the three month period that the assessments had value and that they were useful, mentioning that they were still useful for checking their child's level in class. When asked about whether they were still happy with the new system, the parents also replied positively, with Claire's mother mindful that the assessments had made the institute's overall focus more 'serious' and less 'fun'. As for the overall impact of the system after three months, the parents replied that the new curriculum was a source of motivation for their children to study harder.

However, in contrast with the children's admissions that the amount of pressure they were feeling from the assessments was mounting, the parents in their final interviews downplayed the amount of the stress their children were under. Three of the parents also believed that the assessments were motivating their children to perform well and study harder, with only Ben's mother a little unsure, in contrast with the split responses to this question from the learners. As with the learners however, the parents were beginning to doubt the validity of the tests as the true gauge of their children's English language ability, with Claire's mother mentioning 'Claire likes reading and speaking so I think the tests are not a fair representation' and with Ben's mother going so far as to say 'I don't think they are totally, I wish they would test other things. He really likes reading but there isn't any on these tests'. On the other hand, when asked about the perceived formative benefit of the assessments, the parents were equally unsure as to the usefulness of the washback from the assessments, mentioning:

(CLAIRE'S MOTHER) Teachers don't tell us the other students' test scores, only the average, so it seems like the teachers want to keep that to themselves. They only say

‘your child did well’ and that’s it. They don’t like to tell us why our children got something wrong, we don’t hear anything like that

(JANE’S MOTHER) I didn’t get much, only on the phone. The teacher just said ‘she’s doing fine, her scores are good.’

Claire’s mother’s answer is interesting in that the typical ‘window dressing’ of feedback from teachers to parents that has occurred in this institute (and common with the other institutions I have worked at in my experience) seems to be a double edged sword. On one hand, the teachers cannot usually make particularly critical comments about learners to their parents for fear that the parent will pull their child out the school. On the other hand, parents are not getting the real picture of how their child is performing as any feedback is usually fabricated and therefore lacks much of a useful function.

This lack of post-assessment information carried over onto the progress reports the parents received from the school. Ben’s mother mentions: ‘I can see some good comments on his report, but these things usually only tell us the good news anyway. I have never read anything bad on his report, not just from this school but others as well’ and Suzie’s mother had trouble reading the reports as they were written in English. During my experience working at the institute (as well as the other institutes where I have taught in Korea), teachers were advised to keep student’s reports positive even where student’s performance was particularly poor. As I mentioned previously in this section, this was done to keep the parents ‘happy’ and keep them believing that their student was performing well at the school, so that the parents would not feel the need to pull their student out of the school. This again leads most reports sent home to parents serving a ‘window dressing’

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

function and actually serving little formative (or even truly summative) benefit at all.

Even in terms of whether their child had completed everything in the learning objectives, the parents were unsure about whether the objectives had actually been completed, and believed that most of what was supposed to be achieved in the objectives was to be given as homework, with some of the parents even unsure as to whether the homework was completed or not. As a matter of fact, all of the learners actually mentioned that most of the time they ‘copied’ their homework from another student in the class, or did the homework at their regular elementary school during break time. This could be evidence to suggest that the parents did not appear to know much about the students’ successes or failures regarding the learning objectives in the curriculum.

On a positive note, the new curriculum appeared to have brought with it a new sense of teamwork between parents and teachers, and the parents were happy that they were communicating more with the teachers than they had done previously. This was achieved in part by a new policy put in place by the institute where the native Korean teachers called the parents to discuss the learners’ progress.

On the whole though, as with the learners, the parents were beginning to doubt the effectiveness of the new assessment strategy and seemed unclear as to what kind of formative or summative function the new assessments were providing in practice.

Final Interviews – The Teachers

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

For the final round of interviews, the KETs were very positive as to the benefits of the new assessments. When asked about the value of the assessments over the three month period, both KETs still believed the tests were a useful diagnostic and motivational tool for the learners, with both KETs mentioning that the tests ‘stimulated’ the learners, in contrast with the learners’ admission that the test were a great source of stress. Even if students underperformed on the tests, the KETs felt that the increased sense of competition between the students helped to bootstrap them to a higher level than they would have gained previously, and that students who underperformed would also get more help from their parents. Damien was also very positive as to the impact of the new system after the initial three months, mentioning that the assessments and curriculum appeared to give his classes more ‘structure’ and ‘focus’, and that his students were more motivated to study under the new assessments. Keri, as with her other comments from all of the interviews so far, painted a much more negative picture of the value of the new assessments, with her only positive comment being ‘the good thing is they get to practice more speaking one-on-one with me, and that’s always good. It improves their speaking’.

As for any perceived pressure the learners may have been feeling due to the assessments, both KETs again mentioned that the tests were stimulating for the students and they didn’t seem to feel under too much pressure, and the NETs also downplayed the amount of stress the students were under.

In terms of washback from the tests, the KETs appeared to put the responsibility for these matters on the NETs. This division of responsibility affected a lot of the KETs answers, with Emily especially suggesting that the success or failure of the students as well

as the usefulness and validity of the tests themselves rested squarely on the shoulders of the NETs as the NETs were responsible for writing the questions for the tests. Emily mentions:

(EMILY) Actually, the tests show just whether this student is good or not depending on how the teacher makes the questions [...] if the teacher makes poor questions it won't be a good test. as long as the teacher makes good questions then we can say 'he is good' 'he is bad' otherwise we can't really decide whether the student is good or not.

On the other hand, when asked whether there had been any teamwork between teachers and the parents to improve the learners' abilities, the NETs clearly felt that this was the responsibility of the KETs (presumably due to the language barrier).

Damien still believed the tests held a useful formative function, and he claimed that he used the tests results to guide his lessons, offering 'review' classes if the students missed a key structure or provision on the learning objectives. He also believed the tests still offered a useful summative function as well, stating that 'when I look at the grades from testing of a student, I certainly know how well that student is doing in my class'. The answers of Keri tell a quite different story however. In the answers to most of her questions on the washback from the tests, Keri claims that there hasn't been any 'follow-up' to the assessments and therefore the assessments have been a waste of time. She mentions:

KERI 'if there was follow-up I think it would be a good system, but because there isn't any follow up then no, I don't think it's effective. By follow-up I mean if the kids are

not doing well, place them in different class. If the kids can't read move them down or provide services that can help them.'

She also mentioned that the objectives in the curriculum provided no benefit either, as she feels that the learning objectives were far too difficult for her low level class to achieve. She also felt that the reports to the parents were also a waste of time, going so far as to say 'Sometimes I see the kids still have them in their backpacks a week later so I don't even think a lot of the reports are getting home or if the parents see them.' For Keri then, neither the washback from the assessments nor the learning objectives served any formative or summative purpose at all.

With regards to whether the teachers attempted to condition the students to pass the tests, both NETs stated that they did, in keeping with some of the learners' admissions.

(DAMIEN) well, I would always have like, at least one class before the tests I would always have a review period, where I would introduce them to the type of questions that will be on the test, and we will review the material that I knew would be on the tests.

Whether or not this 'review' period actually constituted the kind of test training that would seriously influence the children's performance on the tests is unclear, however.

In general, the answers of Damien and the KETs remained supportive of the new system, whereas Keri's answers stood in marked contrast. During a conversation I had with my supervisor at the institute who helped me with these interviews, I asked her why

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

the KETs seemed to give such a glowing report of the new assessments over the three months, and she responded ‘Well, they are Korean. You know how they are... they just try to remain positive and not rock the boat too much.’ I felt this opinion was interesting given the nature of the KETs answers, and I felt it was possible that perhaps they were unable or unwilling to share some of their more negative responses with me, particularly as they were aware that my supervisor and I had developed the new assessment procedures and did not wish for us to ‘lose face’. Losing ‘face’ is something that is a big part of Korean culture, due to the presence of ‘기분’ or ‘Kibun’, which translates roughly as ‘mood’, a mood that is strongly related to social harmony. If someone (particularly a foreigner) attempts to disrupt this social harmony in any way, it causes the Korean in question to lose ‘face’ (Kovacs, 2006:http://outsideinkorea.com/inside/2006/06/on_kibun.php). It is certainly possible that the Korean English teachers’ answers may have very well been influenced by ‘Kibun’, and it would perhaps be impossible to really get a true idea about their inner-most feelings regarding the assessments. ‘Kibun’ aside, the answers of the stakeholders over the three month period revealed a lot about both the dangers inherent in assessment in EFL, the most pertinent of which will be discussed below.

Discussion

To review, despite the initial positive perceptions of the assessment process from the stakeholders in the first round of interviews, the stakeholders’ perceptions became increasingly negative in tone during the second and final interviews. To summarize, the main findings of the case studies were that:

- 1) Students felt increased pressure due to the assessments over time, mainly from their
- Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

parents, as they followed the course

- 2) Teamwork and feedback between teachers and parents regarding the assessments and curriculum was not as useful as it was initially thought it might be (especially between the NETs and the parents), although it had got the KETs and the parents communicating a little more due to a new policy of giving feedback over the telephone.
- 3) The assessments themselves did not serve much of a useful summative or formative function (except for Damien) according to the stakeholders, except in the area of rote vocabulary memorization.
- 4) The reports given after the three month period of the case studies lacked much of a formative function and were mostly 'window dressing' in the eyes of most the stakeholders
- 5) The feedback from the assessments for the learners was mostly negative in the form of punishment for poor performance
- 6) There was evidence from the teachers and the learners that training for the assessments had occurred at some level which casted doubt on the tests' real formative validity.

The context in which the case studies were undertaken is almost identical in terms of other institutes and public schools that I have worked at during my work experience so far in Korea.

Limitations

The case studies were severely limited by the language barrier. I was present
Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

throughout each interview, but as I could only really get to the meat of the stakeholder's opinions after the interviews were translated and transcribed, I missed many opportunities to ask further questions pertinent to their answers at the time, which I feel was a big shortcoming of this research, but not one that would invalidate it entirely. Perhaps if I were to attempt this kind of case study in this context again, I would have the translator directly translate to me the stakeholder's answers as they were given, but this would have required more time, and could have in my opinion been seen as a little confrontational by the parents especially keeping in mind '기분' (Kibun).

Implications of the case study

I found that most of the participants of the case study changed their attitudes significantly over the three month duration. The learners and parents had come to the realization that the assessments were not as useful in guiding their learning or assessing their abilities as they believed at first, which I believed to be a serious let-down of the program. The pressure on the learners to perform did build over time, despite their parents' insistence that the tests stimulated the children, and this pressure was damaging to their development in my opinion. The learning objectives were either too difficult or not followed properly so the learners and parents held little stock in them by the end of the case study. Generally I found that the introduction of the new program had negatively affected the learner's attitudes towards assessment and their overall performance and happiness at the school.

In terms of the literature reviewed for the research, particularly our hope to enjoy the results of Brindley's (1994) findings on 'task-centered assessment' when developing Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

the curriculum, in hindsight, only the objectives themselves constituted ‘tasks’, and the monthly assessments were not ‘tasks’ as defined by Ellis (2003), in that the tests were not a workplan (a plan for learner activity), and did not allow users much room (at lower levels of the curriculum at least) to use language pragmatically – rather the students ‘displayed’ language instead (Ellis, 2003:9). Also, in terms of our hopes that the tests would achieve a positive washback effect, with the results of the assessment to be integrated into the learning process, our case study echoes Rea-Dickens and Scott (2007) comments that often washback is ‘unpredictable, depending as it does on many factors, including teacher behaviours, but there is the potential for considerable impact in terms of affective factors.’ (2007:3).

In addition, the results of the case study were consistent with a number of the other findings, namely that high stakes tests increase test anxiety (Benmansour, 1999; Leonard & Davey, 2001; Pollard *et al.*, 2000) and increase the pressure on students to do well (Davies & Brember, 1998; Leonard & Davey, 2001). The students’ answers on both stress and anxiety, particularly in the final month, were also telling in that these students are generally considered to be some of the higher achieving students at the institute, so it would be interesting to compare their answers with some of the students who have a history of underperformance at the institute.

Also, as seen with Damien and his adoption of test-taking strategy classes, teaching became focused on the content of the tests (as with Johnston & McClune, 2000). The true extent to which Damien did this is not known, and it is also unclear as to Damien’s motivations behind his teaching focus. It is possible that he was using the tests in the way

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

they were intended (to have a summative and formative benefit), yet it is also possible that he just wanted his students to get a good score. I believe however, given the positive nature of most of his answers during the case studies, that he believed the tests were serving a useful purpose and for him to give 'review' classes and to concentrate on certain test-related information from the textbook was a perfectly acceptable strategy. However, the focus on the content of the tests could be said to have lead to students adopting test-taking strategies that may have been harmful to learning development (as with Paris *et al.*, 1991; Reay & Wiliam, 1999) in that they began to focus solely on vocabulary at the expense of other forms. In fact, the learners had mentioned that the only benefit of the assessments had been to improve their vocabulary, and the parents had complained that their children enjoyed other areas of study such as reading or making conversation.

From the comments of the parents and the children regarding the meaning of the test scores, there was also evidence of what Torrance (2007) calls the 'displacement' of learning (i.e. understanding) by procedural compliance: i.e. achievement without understanding' (Torrance, 2007:293). I believe that is because the test scores were ceasing to serve a formative or summative function for all involved (with the exception of Damien and the fact that the parents and teachers found the assessments to be motivational). The provisions of the criterion-based curriculum were in Keri's case unachievable, yet every student had to 'pass' them ('there's no follow-up'). The parents were also unclear as to the meaning of criterion-based curriculum, believing after three months that it was meant to guide the students' homework activities (activities which, by the students' admission, were regularly copied from other students).), and the KETs felt the provisions and speaking tests were the NET's responsibility. The parents and teachers' answers also

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

mirrored the findings of Deforges, Holden and Hughes (1994) regarding the substantial variation in parents' views concerning the assessments as well as the relationship between teachers and parents, and despite the increased 'teamwork' between the KETs and the parents at the end of the case study period, there was still little sign that the assessment process had resulted in any major shift of perception regarding the stakeholder's perceived roles on either side.

I found the problems encountered by all the stakeholders in the assessments to be very serious, in that the entire assessment process was unfair and unclear to most people involved, achievers and underachievers included. Those who did well were praised and those who did badly were punished under this flawed system, and I do feel sorry for those learners and the parents of those learners who were supposedly underperforming when they may very well have been progressing satisfactorily. In my opinion, therefore, the tests and criterion-based curriculum in place at the institute had outlived their usefulness and both the learners and the parents at the institute would be better served by assessment of a different format, and some of the possibilities for new assessment formats at the institute are described below.

Alternative forms of assessment / implications for further study.

Looking at possible alternatives, Harlen and Crick (2003) provide a list of positive approaches that may remedy the problems found in the case studies. These approaches include putting less emphasis on grades, 'explaining the reasons for, and the implications of, tests' and 'broadening the range of information used in assessing the attainment of individual students' (Harlen and Crick (2003:201-202).

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

They also mention the following useful strategies when designing assessment procedures, including ‘ensuring that the demands of the tests are consistent with the expectations of teachers and the capabilities of the students, involving students in decisions about testing and developing students’ self-assessment skills and use of learning rather than performance criteria.’ Finally, teachers need to create a classroom environment that ‘promotes self-regulated learning’ (Harlen and Crick (2003:201-202)).

In my opinion, a move to student-regulated assessment would be a very interesting and dynamic experiment on the context in question for this research. The difficulty would lie in trying to make the students overcome the idea of *기분* (Kibun, or social harmony). Getting the students to admit their own shortcomings and give true self assessment would in my view be very difficult, especially considering the fiercely competitive nature of Korean students, schools and parents.

Another approach to assessment that could be beneficial in this context was summarized in Harlen and James (1997), namely that assessment should promote ‘real’ or ‘deep’ learning as opposed to ‘surface’ learning. Deep learning is defined as an intention to provide more personal understanding and for learners to have an active with the content that they are being assessed on. Surface learning, on the other hand, is categorized as the ability to produce content only when required, a passive acceptance of content, and learning is only focused on meeting the assessment criteria. I would suggest that in the institute’s context, and from the results of our case study, that the type of learning that occurs from the assessments is that of surface learning. In order for deep learning to have taken place, a

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

more personal approach must be developed that builds on their existing skills and ideas and is 'owned' by the learner, and are seen by learners as relevant for purposes other than of passing the assessments. Harlen and James suggest 'assessment for this purpose is part of teaching. Learning with understanding depends on it.'(Harlen and James, 1997:369). In my opinion, a more personalized approach to formative assessment would be very beneficial in the context of the case studies and that would mean that the criteria for assessment would have to be re-imagined in terms of each student and class instead of a cohort or 'class level'.

Harlen and James suggest that 'all formative assessment is, by definition, criterion-based' (Harlen and James, 1997:370). However, with a personal approach to assessment, the assessment criteria will be based on a judgment of the student's performance rather than on a set of complex assessment criteria. Learners can then better understand the assessment criteria 'through experiences designed to involve them in looking critically at their own work.' (Harlen and James: 1997:372). This view is supported in Black and Wiliam (1998), who in their case study on the benefits of formative assessment, found that 'formative assessment builds students' skills for learning to learn' (Black and Wiliam, 1998:13). Elwood (2006) builds on this further by recommending that schools need to be developing this approach to formative assessment within teacher education and development programmes as well as school wide and nation-wide assessment policy reviews with a view to letting this approach 'become embedded into the life and work of teachers and students in schools' (Elwood, 2006:220).

The same personalized approach to assessment can be used for summative

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

assessment purposes as well, as long as summative assessment becomes an ongoing, continuous process that does not start or stop on test day. In order for summative assessment to be a true representation of a student's ability and to avoid the often crippling affective factors that may influence formal tests (such as the ones found in our case study regarding stress or other stakeholders' influence) it is necessary to take a more holistic approach to summative assessment. This could mean looking at every piece of a student's work or utterance in class and deciding whether, as a portfolio of sorts, their work over the assessment period 'fits' the criteria for assessment. In this way, students would not feel undue stress or pressure from formal 'tests', and the grading of a student's performance or mastery of the assessment criteria would not be based on a one-off instance but over the assessment period as a whole. To summarize, learning goals are needed, but it is the tracking of the individual student's progress towards those goals that is important, rather than meeting the goals directly. In my opinion, this approach would be possible at the institute found in the case study, perhaps in the form of a portfolio of work completed (written or reading passages) and the lessons being recorded over a period of time for evaluation near the end of the assessment period. I would very much like the opportunity to implement this kind of assessment policy in the future.

Finally, I also find Sadler's (2007) views on accurate and valid assessment to be very useful in this kind of situation also. He states:

'For my money, learners can be said to have learned something when three conditions are satisfied. They must be able to do, *on demand*, something they could not do before.

They have to be able to do it *independently of particular others*, those others being

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

primarily the teacher and members of a learning group (if any). And they must be able to do it *well*. (Sadler, 2007:390-391)

Given the opportunity for further study into assessment at a later stage, I would like to research the possible effects of Sadler's method of assessment on the learners at the institute. The removal of any kind of scaffolding from the teachers or parents (or even other students) with regards to the assessments may provide a clearer picture of the learner's true abilities and motivations. Removing the other stakeholders from play would, in my opinion, create a potentially unspoiled landscape for assessment. How this utopia would be achieved, on the other hand, is unclear, and if this research and case studies is anything to go by, the impact of the various stakeholders on assessments is one that is very large and is not easily preventable.

Bibliography

Alderson, J C and Banerjee, J (2001), 'Language testing and assessment (Part 2)', *Language Teaching* (2002), 35: 79-113 Cambridge University Press

Bachman, L, F (1990) '*Fundamental Considerations in Language Testing*', Oxford University Press, UK.

Benmansour, N. (1999) Motivational Orientations, Self-efficacy, Anxiety and Strategy Use in Learning High school Mathematics in Morocco, *Mediterranean Journal of Educational Studies*, 4, pp. 1–15.

Black, P. & Wiliam D. (1998) Assessment and classroom learning, *Assessment in Education*, 5(1), 7–71.

Brindley, G (1994) 'Task-Centered Assessment in Language Learning: The Promise and the Challenge' in Bird, N, Falvey, P, Tsui, A, Allison, D and McNeill, A (eds.) '*Language and Learning*' Hong Kong Institute of Language and Education, Hong Kong

Broadfoot, P., Dockrell, B., Gipps, C., Harlen, W. & Nuttall, D. (1992) *Policy Issues in National Assessment:*

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

BERA dialogues no. 7. (Clevedon, Multilingual Matters Ltd).

Canale, M and Swain, M (1980) 'Theoretical Bases of Communicative Approaches to Second Language Testing', *Applied Linguistics* 1:1-47

Clarke, C. (1997) 'The Impact of National Curriculum Statutory Testing at Key Stages 1 and 2 on Teaching and Learning and the Curriculum', *British Journal of Curriculum and Assessment*, 7(1),12–18.

Davies, J. and Brember, I. (1998) National Curriculum Testing and Self-Esteem in Year 2 The First Five Years: A Cross-Sectional Study, *Educational Psychology*, 18, pp. 365–375.

Davies, J. and Brember, I. (1999) Reading and Mathematics Attainments and Self-Esteem in Years 2 and 6: an Eight Year Cross-Sectional Study, *Educational Studies*, 25, pp. 145–157.

Desforges, C, Holden, C and Hughes, M (1994) 'Assessment at Key Stage One: its Effects on Parents, Teachers and Classroom Practice', *Research Papers in Education*, 9:2, 133 – 158

Ellis, R (2003) '*Task Based Language Learning and Teaching*', Oxford University Press, UK.

Elwood, J (2006) 'Formative assessment: possibilities, boundaries and limitations', *Assessment in Education: Principles, Policy & Practice*, 13:2, 215 - 232

Evans, E and Engelberg, R (1988) 'Student Perceptions of school Grading', *Journal of Research and Development in Education*, v21 n2 p45-54

Harlen, W and Crick, R D (2003) '*Testing and Motivation for Learning*', *Assessment in Education: Principles, Policy & Practice*, 10:2, 169 – 207

Harlen, W and James, M (1997) 'Assessment and Learning: differences and relationships between formative and summative assessment', *Assessment in Education: Principles, Policy & Practice*, 4:3, 365 - 379

Hoekje, B and Linnell, K (1994) "Authenticity" in Language Testing: Evaluating Spoken Language Tests for

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

International Teaching Assistants', *TESOL Quarterly*, Vol 28, No 1.

Hoekje, B., & Godfrey, D. (1991). *Native speaker/non-native speaker performance and the TSE/SPEAK*. Unpublished manuscript, Drexel University, Philadelphia

Johnston, J. and McClune, W. (2000) *Selection Project SEL 5.1: Pupil Motivation and Attitudes—Self-Esteem, Locus of Control, Learning Disposition and the Impact of Selection on Teaching and Learning*, in 'The Effects of the Selective System of Secondary Education in Northern Ireland', Research Papers Volume II, Bangor, Co. Down, Department of Education

Kirkland, M. C. (1971) 'The Effect of Tests on Students and Schools', *Review of Educational Research*, 41(4), 303–350.

Leonard, M. and Davey, C. (2001) *Thoughts on the 11 Plus*, paper presented at a conference of the Save the Children Fund, Belfast,).

Long, M (1985) 'A Role for Instruction in Second Language Acquisition: Task-Based Language Teaching' in Hyltenstam, K and Pienemann, M (eds.) 'Modelling and Assessing Second Language Acquisition', Clevedon: Multilingual Matters.

McNeil, L. and Valenzuela, A. (1998) *The Harmful Effects of the TAAS System of Testing in Texas: Beneath the Accountability Rhetoric*, High Stakes K-12 Testing Conference, Harvard University, 4 December, 1998.

Paris, S., Lawton, T., Turner, J. & Roth, J. (1991) A Developmental Perspective on Standardised Achievement Testing, *Educational Researcher*, 20, pp. 12–20.

Pollard A., Triggs, P., Broadfoot, P., McNess, E. and Osborn, M. (2000) *What Pupils Say: Changing Policy and Practice in Primary Education*, London, Continuum.

Rea-Dickins, P. (2000). 'Classroom Assessment', in Hedge, T. 'Teaching and Learning in the Language Classroom'. Oxford University Press, UK.

Rea-Dickins, P and Scott, C (2007) 'Washback from Language Tests on Teaching, Learning and Policy: Evidence From Diverse Settings', *Assessment in Education: Principles, Policy & Practice*, 14:1, 1 - 7

Reay, D. and Wiliam, D. (1999) 'I'll be a Nothing': Structure, Agency and the Construction of Identity
Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

Through Assessment, *British Educational Research Journal*, 25, pp. 343–354.

Sadler, D R (2005) 'Interpretations of Criteria-Based Assessment and Grading in Higher Education', *Assessment & Evaluation in Higher Education*, Volume 30, Number 2, April 2005 , pp. 175-194(20)

Sadler, D R (2007) 'Perils in the Meticulous Specification of Goals and Assessment Criteria', *Assessment in Education: Principles, Policy & Practice*, 14:3, 387 – 392

Scott, C (2007) 'Stakeholder perceptions of test impact', *Assessment in Education: Principles, Policy & Practice*, 14:1, 27 – 49

Shepard, L. A. & Bliem, C. L. (1993) *Parent Opinions About Standardized Tests, Teacher's Information and Performance Assessments: A Case Study of the Effects of Alternative Assessment in Instruction, Student Learning and Accountability Practices*. CSE Technical Report 367, University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards and Student Testing.

Swales, J (1990) '*Genre Analysis*', Cambridge, Cambridge University Press.

Torrance, H (2007) 'Assessment as Learning? How the Use of Explicit Learning Objectives, Assessment Criteria and Feedback in Post-Secondary Education and Training Can Come to Dominate Learning. ', *Assessment in Education: Principles, Policy & Practice*, 14:3, 281- 294

Web Sources

Kovacs, C, (2002) '*On '기분' (Kibun)*' Retrieved July 20th 2008 from http://outsideinkorea.com/inside/2006/06/on_kibun.php

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.



PKNU, FLEC
English Conversation Course

Student Identification			
Name	김다수님	Class	K1
		Month	April, 2008
Textbook	Treasure 1.5		
	Sub material		
Week	Objective		
1st Week	Students should be able to talk about what they have to do every day (I have to / I'm supposed to - using the example on Olivia P.26) Students should write a weekly chore plan showing what they do around the home. Students should also read up to page 27 of Olivia		
2nd Week	Students should be encouraged to draw freely on paper using a variety of materials and then use descriptive language to explain what it is they are drawing by using Olivia P.30 as an example. They should then complete the story of Olivia and do the comprehension check on P.37		
3rd Week	Students should complete the section about 'cats in art' and talk about famous cats they like. Students should then write a similar invitation to the one on page 42, using the directions on page 43. Students should be aware of the use of adjectives to make descriptions.		
4th Week	Students should make sentences using the words on Page 46. Then students should read up to page 55 of the story 'The Kite', noting the process of flying a kite. Students should be especially aware of the phrases 'give up' and 'try again'.		

Appendix 1 – Sample of Learning Objectives

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.



PKNU, ELEC
Junior English Intensive Course

Speaking Test

Student Identification

Name	SA	Class	Junior 1	Total Score	/ 100	Class Average	87 / 100
------	----	-------	----------	-------------	-------	---------------	----------

Coursebook	Treasure 2.2	Storybook	N/A	Test Date (yy/mm/dd)	21/05/2008
------------	--------------	-----------	-----	----------------------	------------

Questions

1	What's the first step in recycling?	1	2	3	4	5	6	7	8	9	10
2	What's the second step in recycling?	1	2	3	4	5	6	7	8	9	10
3	What's the third step in recycling?	1	2	3	4	5	6	7	8	9	10
4	What's the fourth step in recycling?	1	2	3	4	5	6	7	8	9	10
5	What's the fifth step in recycling?	1	2	3	4	5	6	7	8	9	10
6	What is a big problem in the world today?	1	2	3	4	5	6	7	8	9	10
7	What is one way we can help the earth?	1	2	3	4	5	6	7	8	9	10
8	What is another way we can help the earth?	1	2	3	4	5	6	7	8	9	10
9	What does extinct mean?	1	2	3	4	5	6	7	8	9	10
10	Which species of animals are now extinct?	1	2	3	4	5	6	7	8	9	10

Instructor's Comment

SA is a nice student and is improving quickly.
 She always tries hard and is always a joy to have in class.

Instructor's Signature

Academic Director

TA



Appendix 2- Sample test sheet and report card
First round (March 2008)

Learners

Do you think the speaking tests are important?

In your opinion, what do you think the test scores mean?

Do you know why we introduced the new system?

Would you feel any uneasiness from the new system?

Do you feel any pressure from your mum from the assessments?

Do you understand the aim of the test?

Do you think the assessments are a good idea?

Parents

What do you think the test scores mean?

Why did we introduce the new system?

In your opinion, what becomes of those who succeed and those who fail under this new system?

Can you think of any negative side to these assessments?

What is the role of the parents and teachers for the learner?

Do you know how the assessments are done?

What is more important for you - test scores or teachers comments?

What do you think will be the effect on the learners of the new system?

Do you think the new weekly learning objectives are important?

Do you understand the purpose of the tests?

Do you think do the learners will feel pressure from these assessments?

Do you think that the tests are really necessary?

In our new system, there are levels 1 to 6. Do you understand how these levels are reached and what they entail?

What information do you want to get from these assessments?

What do you think the benefits of the new system will be?

Do you think that the new system will boost teamwork between teachers and parents?

What do you think the effect of the assessments will be on student's motivation?

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

Appendix 3 - Interview Questions

The Teachers

In your opinion what do the grades from testing actually mean?

Why did we introduce the new system?

What becomes of those who succeed and those who failed under these assessments?

Are there any negative aspects to this method of assessment?

Do the students know what the test scores mean?

What are the stakes for the students in your opinion?

What are the stakes for teachers?

Do you believe the test scores are useful for assessing learning?

Is the new system a good idea?

What do teachers think of the parents' role in this new system?

Do you believe that parents are interested in the new system?

Interview Questions Part 2 – April 2008 – Post assessment

Learners

Are you happy with the grades you have received from the tests so far?

Do these grades have any particular meaning for you as a student? Has the grading been clear?

Have you felt any stress or anxiety from the assessments?

Have you been more motivated to study under the new system?

Did the learning objectives help you to learn/study in any way?

What happens if you cannot do the things mentioned in the learning objectives?

After the tests, did your parents or teachers use the scores to guide your learning in any way?

Do you think the tests have been useful or not?

Did your parents tried to influence your performance on these tests?

How did the teachers get you ready for these tests?

Have you learned about any of your weak points from the new system?

Parents

Are you happy with the grades your children have received so far?

Appendix 3 - Interview Questions

Do these grades have any particular meaning for you as a parent? Are the grades clear?

Have your children expressed any stress or anxiety from the assessments?

In your opinion, has your child been more motivated to study under the new system?

What are your opinions on the learning outcomes in the new curriculum? Are they helpful in assessing what has been learned?

In your opinion, are the learning objectives helpful in any way for the learning process? Do they help students to learn?

What is the effect if a student can do the things mentioned in the learning objectives? Do you believe the students have mastered that objective?

What happens if a student cannot do the things mentioned in the learning objectives?

After the tests, do you know if the teacher used the scores to guide the students learning or not? Did you (the parents) use the scores to guide their learning or not?

Do you believe the monthly tests meet the purposes of the curriculum and the learning of English?

Do you believe the teachers tried to influence the students' performance on these tests?

Did you (the parents) try to influence the student's performance on these tests

Have yourselves and the teachers tried to work together to aid student's performance under the new system yet?

Are you more knowledgeable now about the test process?

Teachers

Are you happy with the grades your students have received from the other teachers so far?

Do these grades have any particular meaning for you as a teacher? Has the grading been clear?

Have your students expressed any stress or anxiety from the assessments?

In your opinion, have your students been more motivated to study under the new system?

What are your opinions on the learning outcomes in the new curriculum? Are they helpful in assessing what has been learned?

In your opinion, are the learning objectives helpful in any way for the learning process? Do they help students to learn?

What is the effect if a student can do the things mentioned in the learning objectives? Do you believe the students have mastered that objective?

What happens if a student cannot do the things mentioned in the learning objectives?

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

After the tests, do you know if the parents used the scores to guide the students learning or not? Did you (the teachers) use the scores to guide their learning or not?

Do you believe the monthly tests meet the purposes of the curriculum and the learning of English?

Appendix 3 - Interview Questions

Do you believe the parents tried to influence the students' performance on these tests?

How did you (the teachers) try to influence the student's performance on these tests?

Have yourselves and the parents tried to work together to aid student's performance under the new system yet?

What kind of comments did you write on the student's reports?

Final Interviews (End of May 2008)

The learners

If you underperformed on the tests, what happened? How did you feel if you were in this situation? If you did well, can you imagine how those who didn't do so well might feel?

Did the tests have any value for you over the 3 month period? Did you find the process useful in any way? What did the grades tell you?

Did the teachers or your parents try to 'train' you to take the tests in any way? Did they try to influence your performance on the tests by showing you how to pass them? How did they do that?

Do you feel anxious about taking the tests? Do you feel stressed out from taking them?

**How much pressure (if any) about the tests did you feel from your parents?
Did you feel any pressure from the teachers over the tests?**

Have the tests or the curriculum affected your motivation to learn English?

Did you complete everything mentioned in the learning objectives? What happened if you didn't get everything finished in time?

Do you think the tests are a fair representation of your real English ability or do you think they don't give you the chance to show everything you can do?

What feedback did you get from the teachers after taking the tests? How about feedback from your parents?

How do you feel about what is written on your report? Is it useful? Is it important for you?

The Parents

Did the tests have any value for you over the 3 month period? Did you find the process useful in any way? What did the grades tell you about your child's performance?

If your child underperformed on the tests, what happened? How did you feel if your child were in this situation? If your child did well, can you imagine how those who didn't do so well might feel?

Are you happy with the new system after 3 months? If so, why? If not, why not?

What has been the overall impact of the new system after 3 months? What positive aspects has it
Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

brought? What negative aspects have come with it?

Has the new system promoted teamwork with the school and the parents over the three month period? How much cooperation has there been? Has it been useful in any way?

Appendix 3 - Interview Questions

Did you try to 'train' your child to take the tests in any way? Did you try to influence your child's performance on the tests by showing them how to pass them? How did you prepare your child for the tests?

How much pressure (if any) about the tests did your child feel? How did this affect your child?

Have the tests or the curriculum affected your child's motivation to learn English?

Did your child complete everything mentioned in the learning objectives? What happened if your child didn't get everything finished in time?

Do you think the tests are a fair representation of your child's real English ability or do you think they don't give your child the chance to show everything they can do?

What feedback did you get from the teachers after taking the tests?

How do you feel about what is written on your child's report? Is it useful?

The Teachers

Did the tests have any value for you over the 3 month period? Did you find the process useful in any way? What did the grades tell you about your student's performance?

If any of your students underperformed on the tests, what happened?

Are you happy with the new system after 3 months? If so, why? If not, why not?

What has been the overall impact of the new system after 3 months? What positive aspects has it brought? What negative aspects have come with it?

Has the new system promoted teamwork with the school and the parents over the three month period? How much cooperation has there been? Has it been useful in any way?

Did you try to 'train' your students to take the tests in any way? Did you try to influence your students' performance on the tests by showing them how to pass them? How did you prepare your students for the tests?

How much pressure (if any) about the tests did your students feel? How did this affect your student's behavior or performance??

Have the tests or the curriculum affected your students' motivation to learn English as far as you can tell?

Did your students complete everything mentioned in the learning objectives? What happened if your students didn't get everything finished in time?

Do you think the tests are a fair representation of your students' real English ability or do you think they don't give your students the chance to show everything they can do?

What feedback did you give the students after taking the tests?

How do you feel about what you wrote on your student's report? Is it useful for the learners? What do you think the parents will do now that they know the average grades? Will the parents pay attention to your comments?

Appendix 3 - Interview Questions

(Interviewer) Did the tests have any value for you over the 3 month period? Did you find the process useful in any way? What did the grades tell you about your child's performance?

(Claire's mother) She started studying English before elementary school, but these 3 months have been the most difficult period for both of us, me and SA... this class is at a higher level so the tests seem mandatory

(Interviewer) Yeah

(Claire's mother) But it's not so fun for her anymore.

(Interviewer) So, what did the grades tell you about your child's performance?

(Claire's mother) When we got the scores for the first time, in the past she didn't get many wrong, so this time she was shocked and stressed out.

(Interviewer) They got stressed out?

(Claire's mother) Yeah.

(Interviewer) Ok, next, if your child underperformed on the tests, what happened? How did you feel if your child were in this situation? If your child did well, can you imagine how those who didn't do so well might feel?

(Claire's mother) Now she is one of the students who didn't do so well... at the beginning I talked to the teacher a lot... now I can see she is getting better and better scores

(Interviewer) But that took about three months right?

(Claire's mother) Yeah, so now she knows how to do it, not 100% perfectly, but she is more confident because of it.

(Interviewer) Are you happy with the new system after 3 months? If so, why? If not, why not?

(Claire's mother) In the beginning I think it worked well and it was well received... but I guess SA and other students, couldn't get used to the new system quickly enough... also in the beginning we realized that this new system was very strict...

(Interviewer) It changed suddenly right?

(Claire's mother) Yeah, it changed so suddenly that she lost interest and through attending these interviews I felt that the teacher's attitude also changed. I think that in the beginning the teacher estimated our children's ability too highly

(Interviewer) Yeah, in the beginning the student's felt very awkward so..

(Claire's mother) so maybe later it will be better.

(Interviewer) What has been the overall impact of the new system after 3 months? What positive aspects has

Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

it brought? What negative aspects have come with it?

(Claire's mother) The positive aspect is that she should study English her entire life, and this system helped her to realize that English is not only something fun but it can also be difficult. They should learn to study like this

(Interviewer) Sort of like higher learning, a higher level?

Appendix 4 – May Interview Claire's mother full transcription

(Claire's mother) This process is important [...] The negative aspect is that everything was in too much of a hurry at first.

(Interviewer) OK, has the new system promoted teamwork with the school and the parents over the three month period? How much cooperation has there been? Has it been useful in any way?

(Claire's mother) Do you mean at school or here at the institute?

(Interviewer) I mean here at the institute

(Claire's mother), Well, in the beginning and even last year I just checked the mail they sent to me, but now there are interviews, as well as telephone conversations with the teachers, so I feel more involved.

(Interviewer) That's right

(Claire's mother) So it's been useful...I can see how well my daughter studies.

(Interviewer) Ok then, another question. Did you try to 'train' your child to take the tests in any way? Did you try to influence your child's performance on the tests by showing them how to pass them? How did you prepare your child for the tests?

(Claire's mother) We tried to go through the words in the book together, but its hard for me too as I have a lot to do and can't spend much time.

(Interviewer) How much pressure (if any) about the tests did your child feel? How did this affect your child?

(Claire's mother) I don't think she feels that much now, but there is still something I can feel.

(Interviewer) Have the tests or the curriculum affected your child's motivation to learn English?

(Claire's mother) It is definitely motivating them now... In the beginning she didn't feel like studying, she asked 'why should I study this?' and said it was difficult to do but now she seems to realize that she should do well

(Interviewer) Yeah, they used to do only conversation so they found it difficult but nowadays they seem OK.

(Claire's mother) Her friends seem to be the same way.

(Interviewer) Did your child complete everything mentioned in the learning objectives? What happened if your child didn't get everything finished in time?

(Claire's mother) Overall I think they did, I can see if they have done them. I check once every two or four days...

(Interviewer) Do you think the tests are a fair representation of your child's real English ability or do you think they don't give your child the chance to show everything they can do?

(Claire's mother) I really cared about the tests a lot at the beginning ...but the tests just seem to be vocabulary
Crosthwaite, P. (2009). The Effects of Assessment on Stakeholders within an Elementary-Grade EFL Program. *Asian EFL Journal*. M.A. Thesis.

memory tests. SA likes reading and speaking so I think the tests are not a fair representation.

(Interviewer) What feedback did you get from the teachers after taking the tests?

(Claire's mother) The teachers don't tell us the other student's test scores, only the average, so it seems like the teachers want to keep that to themselves. They only say 'your child did well' and that's it. They don't like to tell us why our children got something wrong, we don't hear anything like that

Appendix 4 – May Interview Claire's mother full transcription

(Interviewer) OK, last one. How do you feel about what is written on your child's report? Is it useful?

(Claire's mother) Yeah, the reports have a big effect on the kids because they can see each others scores when they give out the papers in class.

Appendix 4 – May Interview Claire’s mother full transcription