

AN INVESTIGATION OF THE CONFORMITY BETWEEN AN ASSESSMENT POLICY DOCUMENT AND IMPLEMENTATION

Emad A.S. Abu-Ayyash

Ministry of Education, United Arab Emirates.

emadayyash@gmail.com

ABSTRACT

This paper explores the extent to which the English Department adheres to the assessment guidelines depicted in the assessment policy document in one of the training centres in the United Arab Emirates. To this end, the study attempts to answer four questions: 1) What are the main components of the Training Centre's Assessment Policy Document?, 2) What are the main guidelines of the assessment practice in the English Department?, 3) What are the differences between the policy's text and implementation? and 4) How can this gap be probably bridged?. In order to conduct the gap analysis, three steps have been made. Firstly, the assessment policy document has been analyzed, and the various sections that compose it have been outlined. Secondly, three semi-structured interviews were made with the three English mentors in the training centre, who were responsible for the assessment procedures and documentation. The questions of the interviews were built based on the assessment policy document's various sections and on the details of each section. Following the interviews, the data provided in the policy document and those provided by the three English mentors in the interviews were compared to find out whether or not the assessment implementation was in congruence with the assessment policy text. The results of the study have revealed a number of disparities between the assessment policy text and the implementation at different levels. Finally, the paper reflects on the gap between policy and implementation by analyzing the possible causes that have probably led to such gap and by proposing a policy model that may facilitate in narrowing the text-implementation gap in assessment..

Keywords: Assessment policy; text-implementation gap

INTRODUCTION

One of the major hindrances of achieving the desired outcomes of institutions of different kinds has been acknowledged to be the gap between the policy text, or document, and the implementation of what is scripted in it (McGee, 2011). Regardless of context type (educational, commercial, industrial, etc.), when the product does not meet the expectations set by policy makers, the implementation of the policy guidelines can hardly be given the benefit of the doubt. Due to the significance of the conformity between what is written in the policy document and what is done on the ground, and owing to the conspicuous dearth in the studies that address the issue, the present study attempts to investigate whether or not there is a gap between the policy text and its implementation in one of the leading training centres in the United Arab Emirates. Because of the sensitivity of the issue investigated in the present paper, and for the sake of anonymity, the real name of the centre has been replaced with the pseudo name of *Training Centre* (henceforth TC). The document that has been chosen to be investigated is TC's Assessment Policy Document.

TC provides a three-year academic program for local school-leavers in the United Arab Emirates. Students who successfully complete the three-year program can then get jobs with a number of industries based on their specializations. A factor of major importance regarding the benefits students receive after graduation when they join their workplaces is their

achievement in final exams in different subjects, such as Information Technology and English. The measurement of the students' achievement all through the entire program should ideally be based on TC's Assessment Policy Document, which was set by the Assessment Unit in the centre in 2009. The current paper examines TC's Assessment Policy from two dimensions: the policy text and its implementation by the English Department in the centre. The choice of this department is not random since the students' achievement in the English language is a determiner of both their graduation and their employment opportunities. Put differently, high levels of competency in English is one of the major conditions of TC's industry partners to hire its graduates. This dimension adds more value to the analysis of whether or not the assessment policy makers and the implementers of that policy are on the same page. Specifically, the present paper is intended to answer the following questions:

1. What are the main components of the TC's Assessment Policy Document?
2. What are the main guidelines of the assessment practice in the English Department?
3. What are the differences between the policy's text and implementation?
4. How can this gap be probably bridged?

LITERATURE REVIEW

As this paper houses a twin focus on assessment and the gap between a policy text and its implementation, the literature review section has been divided into two parts accordingly. The first part focuses on general assessment issues, and the second surveys a number of studies that have investigated the relationship between policy texts and implementation from different perspectives.

Assessment

This part of the review sheds light on the difference between assessment and tests, the different types of assessment and tests and the major characteristics of each type. To start with, Brown (2007) provides a stringent clarification that assessment is not and should not be restricted to tests. He defines a test as "a method of measuring a person's ability or knowledge in a given domain" (p. 445). In his differentiation between assessment and tests, Brown stresses that assessment is an ongoing process that houses a broader domain than tests. In this regard, tests are only subsets, or the premier form of assessment (Poehner, 2008), but they are not what assessment is all about.

A host of authors refer to two major types of assessment: summative and formative (Chappuis & Chappuis, 2008; Garrison & Ehringhaus, 2012; Harlen & James, 2006; Johnson & Jenkins, 2003), only to name some. Table 1 below outlines the various characteristics of both summative and formative assessments as described by different authors.

It is immediately clear that the below accounts harbour patent commonalities that are worth pointing out. As far as Formative Assessment is concerned, it is obvious that the dominant characteristic of this type of assessment is that it is *ongoing*. One more patent feature of Formative Assessment is that it involves students, driven by the teaching-learning process and directed towards monitoring their progress during the instructional process (Miller, 2015; Thum et al., 2015). Summative Assessment, on the other hand, emphasizes achievement and is intended basically to provide some kind of documentation of what students know and do not know. Angelo and Cross (1993) sum up the distinction between the two types of assessment, stating that while Summative Assessment tools are designed to make judgments about individual student achievement and to assign grades, Formative Assessment tools are

designed to inform teaching and improve learning and are used as feedback devices. Although it is not easy to provide an exhaustive list of the types of both kinds of assessment, it does not seem futile to provide a brief account of the main forms involved in Summative Assessment and Formative Assessment.

Table 1. Characteristics of Formative and Summative Assessments

<i>Author (s)</i>	<i>Summative assessment</i>	<i>Formative assessment</i>
Johnson & Jenkins (2003)	<ul style="list-style-type: none"> - assessment of learning - identifies what the student has learned at the end of a unit - takes place at certain intervals 	<ul style="list-style-type: none"> - assessment for learning - monitors students' progress - provides immediate feedback - is directed towards promoting learning
Harlen & James (2006)	<ul style="list-style-type: none"> - relates progress in learning to certain criteria, or goals - is based on evidence - used to measure the learner's level and the program's effectiveness 	<ul style="list-style-type: none"> - considers the learner's progress and effort - requires the involvement of students
Chappuis & Chappuis (2008)	<ul style="list-style-type: none"> - documents how much learning has taken place 	<ul style="list-style-type: none"> - an ongoing, dynamic process - provides information during the teaching and learning process
Garrison & Ehringhaus (2012)	<ul style="list-style-type: none"> - given mainly to decide what students know do not know - given periodically at a particular point in time 	<ul style="list-style-type: none"> - part of the instructional process - provides the data necessary to adjust teaching and learning while they are taking place - involves students in the assessment process

Starting with Summative Assessment, it encompasses different forms based on the choices made by different academic units. One of the most common summative forms of assessment is the end of unit/chapter/semester/placement tests (Coffey, 2012). Akin to this view, Johnson and Jenkins (2003) state that “When assessments reflect the stated learning objectives, a well-designed end of unit test provides teachers with information about individual students (identifying any student who failed to meet objectives), as well as provides an overall indication of classroom instruction”. Therefore, tests primarily link to certain criteria, or objectives, to be met by learners. These types of end-of-course tests are often labelled as *Achievement Tests* (McNamara, 2000). As a major form of Summative Assessment, tests can be classified into different types. *Table 2* summarizes different types of tests as introduced by Brown (2007) and Davies (1990).

Table 2. Types of tests (Brown, 2007; Davies, 1990)

<i>Types of tests</i>	<i>Description</i>
Proficiency test	A test that is not limited a single course or curriculum and that taps global competence in a language, such as IELTS and TOEFL.
Diagnostic test	A test designed to diagnose a particular aspect of language, such as pronunciation, writing, or reading.
Placement test	A test designed to place learners into appropriate levels.
Achievement test	A test that is related to particular content covered in a curriculum in a particular time.

Another major form of Summative Assessment involves the use of portfolios. Portfolios are summative “When used as part of an evaluation of student learning, portfolios provide evidence to support attainment of stated learning objectives” (Johnson and Jenkins, 2003).

Put differently, any form of testing that aims to provide evidence on learners' overall achievement as aligned with certain criteria is considered summative. Formative Assessment takes forms that are different by nature. This difference between formative and summative forms is nested in the very purpose of both types described earlier.

As Formative Assessment is by definition an ongoing process that provides immediate feedback to learners, its types reflect this dynamicity and are, therefore, different from those tools used in Summative Assessment. It might be instructive, however, to start with the reasons why some educational institutions might be opt to use Formative Assessment. Chappuis and Chappuis (2008) provide the following list of advantages of this type of assessment. First, since this assessment type is about immediate feedback, the timeliness of results will make it possible for teachers to instantly adjust their instruction, while learning is in progress. Another advantage of formative assessments is related to the immediate receivers, or the students. By way of explanation, Chappuis and Chappuis (2008) state that the students who are assessed are the ones who benefit from the adjustments. This is not hard to believe given that learners are provided with on-the-spot feedback. Finally, the authors believe that with formative assessment, there is a big opportunity for students to adjust and improve their own learning.

Brown (2007) centres the idea behind formative assessment on performance and refers to this type of assessment as "performance-based assessment". He argues that "instead of offering paper-and-pencil single-answer tests or possibly hundreds of discrete items, performance-based testing of typical school subjects involves open-ended problem solving tasks, hands-on projects, student portfolios, experiments, tasks in various genres of writing and group projects" (p. 460). It should be noted that portfolios are considered as formative rather than summative only when they are used to create a record of a student's development in a number of areas (Johnson & Jenkins, 2003). A more inclusive list of formative types encompasses portfolios, journals, logs, conferences, self-assessment, peer-assessment, self-assessment combined with peer assessment, group work, pair work and group work combined with pair work (Brown, 1998; Garrison & Ehringhaus, 2012).

Policy Text and Implementation

The gap between policy text and implementation has been around for decades now (Haines, Kuruvilla & Borchert, 2004). The following review sheds light on the gap between the written policy text and the practice of policy as viewed and highlighted by different authors and researchers.

Stigler and Hiebert (2004) describe the gap between the policy document and its implementation as a slippage that has resulted because of the lack of a shared language between policy makers and teachers. In congruence with this claim, McLaughlin (1987) argues that the major problem about policy implementation is that by and large it is not easy to make something happen across layers of government and institutions. The author links implementation to two focal factors: capacity and will. While the former is easy to cater for, the latter, represented in motivation, attitudes and beliefs, is "less amenable to policy intervention" (p. 172). According to McLaughlin (1987), what makes *will* a hard nut to crack as far as policy implementation is concerned is that it can be influenced by factors, environmental stability and contending priorities for example, that are hard to be controlled by policy. Alarming, the article describes teachers as "resistant to change", or even lazy implementors. To McLaughlin, the gap between policy and implementation can be better bridged by moving to "a model of social learning and policy analysis that stresses reflection and assistance to on-going decision-making" (p. 175).

Jerald (2005) refers to internal and external factors that challenge implementation in schools. As far as the internal factors are concerned, Jerald identifies technical, cultural and political challenges. While technical challenges involve lack of knowledge about new strategies, cultural challenges involve ingrained behaviour and beliefs that are at conflict with new ideas. Political challenges mainly rise from conflicts among competing interests. Jerald's external factors are very much similar to McLaughlin's concept of capacity. By way of elaboration, Jerald puts external challenges at insufficient support, insufficient control over budgets and insufficient control over personnel. However, unlike McLaughlin, who believes that capacity is easy to deal with, Jerald argues that external factors are severe and very hard to control.

The gap between text and implementation has been characterized by a dispute over a top-down or a bottom-up model. O'Toole (2000, p. 268) states that "an important dispute between advocates of the so-called top-down perspective and those identified as bottom-uppers raged for a number of years, to some limited overall impact". Following are some studies that tackle the issue of text as an "upstairs" document and implementation as something directly related to people "downstairs".

Levinson et al. (2009) propose some ideas as to how to bridge the gap between policy text and implementation. The authors' argument seems to revolve around two major notions, which are *appropriation* and *democratization*. The article seeks to "inspire democratic dialogue, and to foster and empower participatory agency in the democratic production of policy" (p. 770). According to the article, if policy text is to be put into effective motion, the concept of appropriation has to be adopted. The authors define appropriation as "the possible recursive influence of local actors in the formation of authorized policy" (p. 779). It does not seem much to argue that the article's proposal for bridging the gap between text and practice revolves around espousing democratization and appropriation, and, hence, involving nonauthorized policy actors, such as teachers and students, in the making of the policy.

In line with Levinson's concepts of democratization and appropriation, McDonnell (1994) suggests a twin focus on persuasion and regulation of assessment policy. By and large, the article suggests that "the policy uses of assessment might be more effective through a fusion of political and expert perspectives" (p. 396). The reason behind the importance of persuasion in assessment policy lies in what is called *hortatory policies*, where assessment is deeply nested in strong beliefs and values. McDonnell maintains that because hortatory policies, such as assessment, are reliant on beliefs and attitudes, persuasion is the tool that has the upper hand, not rules or money for example.

Apparently, both Levinson (2005) and McDonnell (1994) highlight the indispensable role of people "at the bottom" in making the policy and implementing it. In line with this view, Brighouse and Woods (1999) consider involvement of teachers as a key characteristic of a successful school. If those actors resist a certain policy, it will remain on paper. Morris and Scott (2003) provide a sound example on how futile is any attempt to implement policy if the real actors resist it. The article indicates that teachers' resistance of the educational policy in Hong Kong around 1997 made that policy into something symbolic, and the policy remained at the level of rhetoric. The implementation phase of the educational policy in Hong Kong turned to be "impossible dreams" due to the lack of co-ordination between central policy makers and teachers, or street-level bureaucrats. The way out of this intricate situation, the article suggests, is a bargaining process that takes place between teachers and the "wise men" employed by the government to design educational policies.

METHOD

The present paper utilises two qualitative methods to gather and analyse data: Document analysis and interviews. Quite understandably, “In policy research, almost all likely sources of information, data, and ideas fall into two general types: documents and people” (Bardach,, 2012, p. 83). Therefore, document analysis has been used to cater for the former and interviews for the latter. Using the combination of document analysis and interviews enabled me to provide a rich description of the nature of the gap between the policy text and its implementation. Another merit of combining the two methods stems from the sequential application where document analysis of the policy text was done first to inform the building up of the questions that were used in the interviews.

Document analysis is a significant data collection strategy that focuses on the research question. There is an obvious concern about documents, though. They are independent of the research. In other words, documents have been out there, and it is the duty of the researcher to decide on the parts of the document that are strongly pertinent to his/her research questions. Still, it is acknowledged that “using documentary material as data is not much different from using interviews or observations” (Merriam, 2009, p. 150). Nevertheless, document analysis has been underused in qualitative research due to the difficulty researchers might face in identifying what documents will be useful to answer the research questions on the one hand and in assessing the authenticity of the selected documents (Merriam 2009). These two concerns about document analysis were overcome in the present research as there was very little need to look for documents other than TC’s Assessment Policy, which was provided by the Assessment Unit, thus waving off any concerns about the document’s authenticity.

In order to verify whether the TC’s assessment policy text is implemented by English teachers or not, interviews were held with the three mentors appointed by the English Team Leader to prepare the pacing schedule, write different exams and register the course work marks and the grades of the final exams. Interviews were preferred over the use of questionnaires because whether the assessment policy is implemented or not is an issue that may require lots of clarification. According to Kumar (2011), interviews are more appropriate for complex situations, and are more useful to collect in-depth information. In addition, some questions about the nature of assessment, such as formative and summative, might have needed to be explained to the participants, an opportunity that might not be possible to do if a questionnaire was used. For even more clarification and seeking in-depth information, the interview was an unstructured one, allowing open discussion and more flexible questioning.

The interviews core questions revolved around the components of the assessment policy text and to what extent they are implemented. Therefore, the questions were centred on the types of assessment used by TC’s English Department, what forms of assessment (achievement tests, placement tests, etc.) are used, the learning program distribution (learning cycles and semesters), the grading scheme (course work marks and final exam marks), the final exams criteria and other emerging issues during the interview. Following is a list of the main questions asked in the three interviews:

- How is the academic year divided?
- How is streaming of students done in the department?
- What type of assessment is used in the TC? Is it summative or formative?
- What form of assessment do you have in the TC?

- What grading system do you have in the English Department? What is the breakdown of the course work marks and the final exam?
- Why don't you follow the same breakdown of the policy text? What is the difference between unit tests and quizzes?
- Do you conduct any international exams, such as IELTS, TOEFL etc.? What is it?
- How does the curriculum align with the requirements of the international exam, if there is one used in the TC?
- What points of contact do you have with the Assessment Unit in the TC?
- Do you think the current Assessment Policy in the TC is effective? Why? Why not?

All the three mentors were sent these questions one day before the interview so that they can prepare for the answers. Each interview approximately lasted for 10 minutes, and all of them were recorded.

FINDINGS

The TC's Assessment Policy Text

The TC's policy text starts with a brief introduction that is followed by six sections. By and large, these sections are stated as (1) learning cycle grade calculation, (2) semester grades, (3) assessment, (4) assessment preparation, (5) grading and (6) graduation classification. Finally, the assessment policy document concludes with a sample test from the Automotive stream, or specialization. Since this paper primarily focuses on the English Department, only the policy elements that are related to this department will be taken into account. Following is a rough explanation of the introduction and each one of the six sections.

The policy's introductory part clearly states that the centre's program may last for four years depending on the student's stream. The introduction adds that the two factors which determine a student's stream in the centre are age and ability, which necessarily means that students from different age groups cannot be in the same class or follow the same stream pathway. As far as the academic year is concerned, the policy states that each year is divided into two semesters, and each semester into two learning cycles. *Table 3* below summarises the information stated in the policy's introduction.

Table 3. Assessment Policy - General Information

<i>Duration</i>	Two to four years depending on stream
<i>Stream determinants</i>	Age and ability
<i>Academic year distribution</i>	Two semesters – four learning cycles

The first section of the policy, entitled "Learning Cycle Grade Calculation" identifies the weighting of each assessment mode out of one hundred. The policy places equal weighting of 15 marks on quizzes and tests, participation, homework and projects. The remaining forty marks are allocated for 'end of term' and 'end of year' exams. This part also states that 'homework' is class-based.

The second section clarifies grades' distribution per semester. According to the policy text, in the Foundation Year, sixty marks are allocated for the Course Work (quizzes, participation,

homework and projects) and forty for the End of Term (EoT) written exam. It is also stated that in the first Learning Cycle (LC), there is no final test. Therefore, the average of Course Work Marks of LC1 and LC2 is calculated out of sixty, and then the mark of the EoT Test out of forty is added, producing the mark of the Term 1 out of one hundred. The same is applied on Term 2, and then the average of the two terms will be calculated by adding up the marks of Terms 1 and 2 and dividing them by 2 to get a final mark out of one hundred. As far as Years 2 and 3 are concerned, the same process applied except that the Course Work Mark becomes out of seventy rather than sixty.

Moving to the third section of the policy, it provides some details about the exams and re-sits. One of the main items is setting the pass mark at 50%. Another point is that students should pass all subjects in order to be able to proceed with the program. Linked to this issue is what the policy refers to as ‘Referral’. What is meant here is that a student who only fails one subject (a referral) will be allowed to move to the next level provided that he passes that subject at any point in the following term.

The fifth item in Section 3 identifies a maximum of three trials for students with specific learning needs. In addition, the policy states that delay in submitting reports and assignments will lead to a penalty that does not exceed 10 percent of the overall mark. Another important item states that the progress report students get at the end of each term should include enough details about the student’s abilities, areas of weakness and other guidelines. Finally, the policy states that in the final semester, students go for work placement at the end of which they submit a report.

Probably, the most important section of the Assessment Policy apropos the purposes of the current paper is the fourth section as it accounts for the evaluation criteria. The policy identifies eleven criteria for the final written exams. Table 4 lists ten of these as they appear in the policy text. The one excluded from the list is related to Islamic Studies and Arabic, which is not part of the scope of the current study.

Table 4. Final Exam Criteria

<i>No.</i>	<i>Final Exam Criteria</i>
1.	The questions should assess the learning outcomes of the program.
2.	The questions should cover the curriculum taught in the Learning Cycle.
3.	The questions should be varied so that they assess the students’ various aptitudes and skills.
4.	The questions should cover different student levels, commencing with the easiest.
5.	All questions must be clear and free from any ambiguity.
6.	The time allocated for the examination should be long enough for an average student to answer and review all questions.
7.	Complex questions should be avoided.
8.	The Subject Coordinator in charge will set question papers without stating whether it is a Learning Cycle Test or a re-sit exam.
9.	The Subject Coordinator will provide the corresponding model answers on a separate sheet on which he will indicate the marks allocated against each question and part question.
10.	The questions must be written and printed accurately and clearly so that the students may read them easily without assistance.

With regard to the Course Work Marks, they are determined based on unit tests and quizzes, the number of which is based on the number of classes delivered per week; homework and projects. Projects can be measured by referring to the students' portfolios, which as far as English is concerned, should include assignments, quizzes, field reports, project/task reports, samples of students' day-to-day work, copies of self-assessment and peer assessment instruments and checklists of skills mastered.

Section Five of the Assessment Policy indicates that grading should only be conducted by 'accredited assessors' in possession of the Australian TAFE (Technical and Further Education) Award. Finally, Section Six specifies Grades classification and mainly identifies the pass mark as 50.

TC's Assessment Practice

Based on the interviews with the three mentors, (henceforth, M1, M2 & M3), following is a detailed description of the assessment procedures that are currently followed by the English Department in the TC. To start with the academic year, it is divided into three trimesters, following the new division adopted by the UAE Ministry of Education. Once students arrive to the TC, they are streamed according to their level of English. The streaming itself is based on a placement test (Brown, 2007) that is built on the content students study, which is a series called *Ventures*. M1 elaborated that this placement test was the only measurement based on which students are streamed into different classes.

As far as assessment is concerned, all the three mentors explained that the TC espoused Summative Assessment (Brown, 2007; Davies, 1990). The discussions revealed that the centre relied on achievement tests (Johnson & Jenkins, 2003) and quizzes to assess students' levels. Students only move from one grade level to another when the total of the final exam and the course work marks exceed 50. If students, however, achieve less than 50, they have to take a re-sit exam.

While the final exam gets 40 marks out of hundred, the other 60 are given as course work marks. Mentors send these out to the Assessment Unit without clarifying the breakdown based on which the 60 marks are provided. The breakdown of the course work marks is shown in Table 5, which is based on the three mentors' responses.

Table 5. Breakdown of course work marks in the English Department

<i>Quizzes</i>	<i>Portfolio</i>	<i>Listening</i>	<i>Speaking</i>	<i>Participation</i>	<i>Attendance/behaviour</i>	<i>Total</i>
10	10	10	10	10	10	60

In the graduation year, students sit for an international exam, or a proficiency test (Brown, 2007), which is the Key English Test (KET). According to M2, KET is an international ESOL exam introduced by Cambridge University. When asked about whether it is aligned with the Common European Framework of Reference (CEFR), M2 confirmed that it is and that it equals IELTS band score of 3.0. The pass mark of KET is 70. Students who get 85 and above will get KET with Merit, which is equal to IELTS band score of 3.5. M3 added that students who could achieve KET in their first year, would be given the chance to take the Preliminary English Test (PET), another ESOL examination that is equal to IELTS band score of 4.0. Surprisingly, when the mentors were asked whether the English curriculum is aligned with or prepares students for KET, the mentors answer was an emphatic "No".

The next question raised the issue of communication between mentors, who write all the quizzes and final exams, and the Assessment Unit. All the three mentors confirmed that the

points of contact with the Assessment Unit are limited to certain occasions, such as sending the examinations and course work marks to the assessment unit. Of no surprise came their confirmation that none of them is consulted about the TC's Assessment Policy.

DISCUSSION: THE GAP

There is enough justification why policy alone cannot guarantee the best outcomes. McLaughlin (1987, p. 172) asserts that "the consequences of even the best planned, best supported, and most promising policy initiative depend finally on what happens as individuals throughout the policy system interpret and act on them". In other words, policy is not only about the written text, but it is also about the individuals who will implement what has been written. This section highlights the gap between the TC Assessment Policy and the implementation.

The interviews with the three mentors of the English Department reveal that the TC's policy text houses elements that are not updated in the first place. An example of these is the very division of the academic year. The policy text clearly indicates that the academic year in the TC is divided into four learning cycles that fall within two semesters. All the three mentors, however, confirm that currently the academic year is divided into three trimesters and that the learning cycles system is no longer used. This discrepancy between text and implementation indicates that the current policy text needs to be updated as per the new developments in the division of the academic year.

More alarming is the gap between text and implementation as far as the course work marks are concerned. Obviously implementation is very much different from text in various dimensions. First of all, the breakdown of course work marks implemented by the English teachers divides the sixty marks evenly onto listening, speaking, quizzes, portfolio, participation and attendance/behaviour. This distribution is at odds with the breakdown stated in the policy text, which divides the sixty marks evenly on unit tests, quizzes, homework and projects/portfolios. When asked about the reason why the English Department is using a different course work grading scale, the mentors indicated that listening and speaking are not tested in the final exam, which is why they should be accounted for in the course work marks. Another surprising gap is that while the text policy allocates 15 marks for quizzes and another 15 for unit tests, the mentors stressed that the quizzes they write are the unit tests themselves. Therefore, instead of the 30 marks stated in the policy text, quizzes and unit tests get only 10 marks according to the English Department's breakdown of course work marks. The reason why this system is not creating any problems is probably that the Assessment Unit only collects the marks out of 60, without asking for the breakdown.

The interviews also reveal some kind of shortcoming in the policy text. The text does not account for many kinds of tests that are implemented in the TC. For example, the placement test is not referred to in the text although it is the only criterion used to stream students. Another major element lacking in the policy text is the description of the KET, which is the actual benchmark for all students, who have to sit for this test before they can graduate. These gaps call for some reflection on the current relationship between the Assessment Unit and the implementer, teachers in this regard.

IMPLICATIONS AND RECOMMENDATIONS

Quite understandably, assessment procedures should be at the core of school improvement efforts (Hopkins, 2001). The Assessment Unit in TC is a pivotal department by virtue of its tasks and duties. Probably, the most significant role of the assessment unit lies in that it issues the final reports about students' performance. Therefore, it can be acknowledged that the

Assessment Unit is the real link between students and the TC's industry partners, who ultimately hire the TC's graduates based on their grade reports. The gap between the text of the Assessment Policy and its implementation is no longer enigmatic, and both policy issuers and implementers should not turn a blind eye on the significance of bridging this gap. By way of exemplification, when a certain TC student gets 40 out of 60 in his course work mark, it is assumed by the Assessment Unit that this student has been evaluated based on his achievement in unit tests, quizzes, homework and projects/portfolios. However, this is not true as these four elements compose only 20 marks of the total mark: 10 marks for unit tests and quizzes and 10 marks for portfolios, according to the breakdown scheme of the English Department. It is obvious, then, that the results of students are built on false assumptions as far as the Assessment Unit is concerned.

It is hard to believe that the gap between the policy text and implementation at TC can be narrowed down without building bridges between the people who write the policy and the people who implement it (Levinson, 2009; McDonnell, 1994; McLaughlin, 1987). Apparently, what has created the gap between policy makers and implementers at TC is a mere lack of communication. English teachers have beliefs about how assessment should be conducted and how course work marks should be distributed. Those beliefs are not congruent with the assessment policy guidelines, which is why the implementation does not reflect the text once and for all. In great part, the problem stems from the fact that the assessment policy is a top-down document since policy writers have not consulted teachers in the course work marks breakdown. The other shortcoming about the policy text, which does not account for placement tests and the KET, is a bottom-up one since teachers have not updated people in the Assessment Unit about what they know and do in the English Department.

Having diagnosed the problem as "a lack of communication" issue, it becomes unequivocally justifiable to propose democratization of the assessment procedure as a gap-bridger. As far as the TC's Assessment Unit is concerned, democratization can be simply taken for sharing and involving teachers in setting the policy text. According to Trafford (1997), democracy entails active participation, which in turn requires negotiation and compromise. Nevertheless, this sharing practice should not be limited to a certain text or a certain phase of implementation. In order for democratization to be fruitful, it should be part of the TC's culture and an everlasting procedure, or what Stoll and Fink (1989) refer to as "ongoing discussion and reflection" (p. 70). The importance of continuity stems from the fact that reaching an agreement may not be achieved in one or two sessions of dialogue. Instead, reaching a satisfactory-for-all assessment model may require months, if not years, of discussion, implementation and reflection. In a nutshell, if the gap between the TC's assessment policy text and implementation is to be bridged, all assessment procedures should be negotiated, implemented based on dialogue and reflected on continuously in an ongoing process.

REFERENCES

- [1]. Angelo, T., & Cross, K. (1993). *Classroom assessment techniques*. San Francisco: Jossey-Bass, Inc.
- [2]. Bardach, E. (2012). *A practical guide for policy analysis: The eightfold path to more effective problem solving (4th ed.)*. London: Sage.
- [3]. Brighouse, T. & Woods, D. (1999). *How to improve your school*. London: Routledge.
- [4]. Brown, H. D. (2007). *Teaching by principles*. New York: Pearson Education.
- [5]. Brown, J. D. (1998). *New ways of classroom assessment*. Alexandria: TESOL.

- [6]. Chappuis, S., & Chappuis, J. (2008). The best value in formative assessment. *Educational leadership*, 65 (4), 14–19.
- [7]. Coffey, H. (2012). *Summative assessment*. Learn NC. [Accessed 14 September 2015]. Retrieved from <http://www.learnnc.org/lp/pages/5233>.
- [8]. Davies, A. (1990). *Principles of language testing*. Oxford: Basil Blackwell Ltd.
- [9]. Garrison, C., & Ehringhaus, M. (2012). *Formative and summative assessments in the classroom*. Dover: Measured Progress.
- [10]. Haines, A. Kuruvilla, Sh., & Borchert, M. (2004). Bridging the implementation gap between knowledge and action for health. *Bulletin of World Health Organization*, 82(10), 724-733.
- [11]. Harlen, W., & James, M. (2006). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365–379. doi: 10.1080/0969594970040304
- [12]. Hopkins, D. (2001). *School improvement for real*. London: Routledge.
- [13]. Jerald, C. (2005). *The implementation trap: helping schools overcome barriers to change*. Washington, DC: The Centre for Comprehensive School Reform and Improvement.
- [14]. Johnson, E., & Jenkins, J. (2003). Formative and summative assessment. *Education.com* [online]. [Accessed 20 May 2012]. Retrieved from <http://www.education.com/reference/article/formative-and-summative-assessment/>
- [15]. Kumar, R. (2011). *Research methodology* 3rd ed. London: SAGE Publication Ltd.
- [16]. Levinson, B. et al., (2009). Education policy as a practice of power: theoretical tools, ethnographic methods, democratic options. *Educational Policy*, 23(6), 767–795. doi: 10.1177/0895904808320676
- [17]. McDonnell, L. (1994). Assessment policy as persuasion and regulation. *American Journal of Education*, 102(4), 394–420. Retrieved from <http://www.jstor.org/stable/1085462>
- [18]. McGee, B. (2011). The gap between policy and implementation. *CISCO*. [Accessed 4 December 2015]. Retrieved from <http://blogs.cisco.com/security/the-gap-between-policy-and-implementation>.
- [19]. McLaughlin, M. (1987). Learning from experience: lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9(2), 171–178. doi: 10.3102/01623737009002171
- [20]. McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- [21]. Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass.
- [22]. Miller, A. (2015). Formative assessment is transformational. *Edutopia*. [Accessed 10 November 2015]. Retrieved from <http://www.edutopia.org/blog/formative-assessment-is-transformational-andrew-miller>.
- [23]. Morris, P., & Scott, I. (2003). Educational reform and policy implementation in Hong Kong. *Journal of Education Policy*, 18(1), 71–84. doi: 10.1080/0268093032000042218

- [24]. O'Toole Jr, L. (2000). Research on policy implementation: assessment and prospects. *Journal of Public Administration Research and Theory*, 10(2), 263–288.
- [25]. Poehner, M. (2008). *Dynamic assessment*. Pennsylvania: Springer Science.
- [26]. Stigler, J.W., & Hiebert, J. (2004). Improving Mathematics teaching. *Educational Leadership*, 61(5), 12-17.
- [27]. Stoll, L., & Fink, D. (1989). *Changing our schools*. Berkshire: Open University Press.
- [28]. Thum, Y. M., Tarasawa, B., Hegedus, A., Yun, X., & Bowe, B. (2015). Keeping learning on track: a case study of formative assessment practice and its impact on learning in Meridian School District. *Northwest Evaluation Association*. [Accessed 13 December 2015]. Retrieved from https://www.nwea.org/content/uploads/2015/04/Keeping-Learning-on-Track-A-Case-study-of-Formative-Assessment-Practice-and-its-Impact-on-Learning-in-Meridian-School-District_APR_2015.pdf
- [29]. Trafford, B. (1997). *Participation, power-sharing and school improvement*. Nottingham: Educational Heretics Press.