TEACHERS' BELIEFS AND PRACTICES ABOUT ESP ASSESSMENT: A PHENOMENOGRAPHIC STUDY

Chiung-Wen Chang

R.O.C. Military Academy, Taiwan
E-Mail: chiungwenchang@gmail.com

Abstract. *This paper outlines the qualitative research method and results of an investigation of the beliefs about ESP assessment held by six teachers of ESP courses at four universities in Taiwan. Analysis of these interviews indicated that the orientations that teachers adopt to conceptualize curricula for ESP instruction appear to influence the ways in which they assess their students' achievement. This distinction was especially prominent with respect to the methods of assessment the teachers said they utilized and the types of achievements in ESP courses they perceived their students as making. The subject teachers took a Hymesian view of communicative competence by considering language proficiency in terms of real world criteria which, as subject teachers, they felt well-qualified to assess. The English teachers on the other hand veered towards the 'weaker' (and arguably more conservative) approach by focusing more closely on what they are trained to assess, that is, the quality of the language sample elicited through the teaching task. The analysis illustrates how teachers' unique positions, experiences and perspectives affect their implementation of ESP assessment within their classrooms.*

Key words: *ESP, teacher's beliefs, classroom assessment, phenomenographic study*

1. INTRODUCTION

Globalization and the rapid international development of science and technology are encouraging the mobility of students across the English-speaking world and beyond. In order to be able to take up study or work opportunities, knowledge of a foreign language is essential. At present, it is increasingly important not only to be able to use a foreign language, but also to be able to demonstrate that one can use it at the level required by employers, schools, or universities. Considering this, English for Specific Purposes (ESP) assessments are gaining increasing attention, and are becoming an indispensable tool in the modern education system.

In Taiwan, ESP courses which were initiated to teach English based on learners' needs often result in teachers having to cover a wide range of knowledge domains and to accommodate individual learning differences. Moreover, the lack of standards for all ESP courses has complicated ESP teaching. The lack of available ESP teachers presents another significant challenge for ESP programs. Qualified ESP teachers are scarce because most English teachers lack training in specific professional fields (Tsou 2009). Hutchinson and Waters (1987:158) describe ESP teachers as "reluctant dwellers in a strange and uncharted land," a land that Belcher (2006) says is, for many, intimidating. Testing and evaluation for ESP should be carried out in accordance with teaching contents and objectives; the role of testing is to evaluate students' professional knowledge and linguistic competence. Special

stylistic features relevant to the domain, and the application range of ESP should also be considered. Different ESP courses and skills demand various testing approaches. Many of the ESP instructors are not domain experts, which makes the assessment task all the more daunting. Further, since the advent of ESP programs, how classroom assessment has been conceptualized by ESP and subject teachers remains unclear. Thus far, several studies have looked into teachers' assessment practices, with a few focusing on teachers' beliefs and knowledge affecting their decision-making processes in classroom-based assessment (Chang 2005; Davison 2004; Yin 2005). No research of this kind, however, has been done in the context of ESP courses.

The present study addresses this issue by exploring the knowledge and practice of teacher assessment through a qualitative inquiry into the experiences of six ESP teachers at four universities in Taiwan. The objective of the inquiry is to illuminate how teachers perceive and implement ESP classroom assessment. In particular, this study focuses on the identification of teachers' beliefs that likely ground their ESP assessment decisions and practices.

This study is an attempt to assist those officials responsible for providing support to ESP teachers. It will provide a framework for a novel, more needs-based assessment training curriculum - one that is rooted in teachers' ESP classroom experiences and that takes into account the issues that teachers actually deal with in the course of their work. In addition, an insight into teachers' actual experiences in assessment practices will be of help in providing detailed instructions to teachers who are carrying out ESP classroom assessment for the first time, as well as offering a comparative case study for those who have already conducted assessment in their own classrooms.

2. LITERATURE REVIEW

This section offers an overview of the literature informing this study. It begins with a review of the literature on teachers' beliefs and practices. Then the notion of classroom assessment is introduced. Finally, the literature on ESP testing is reviewed.

2.1. Teachers' beliefs and practices

This study is concerned with teachers' beliefs and practices regarding ESP assessment. The mental 'steps' a teacher takes when implementing ESP assessment are largely unexplored. In some recent studies, references have been made to these thought processes within the framework of teacher cognition (Woods 1996). Based on the premise that the beliefs of teachers regarding their professional roles are closely linked to the learning and achievement of their students, the focus of this part aims to review these studies that are concerned with the relationships between teachers' beliefs and teaching practices.

2.1.1. Definitions of teachers' beliefs

Research on teachers' beliefs began to find its way into the field of language teacher education in the early 1990s (Binnie-Smith 1996; Borg 2003). The belief which is under discussion has been variously termed "teacher knowledge" (Freeman 2002), "teachers' mental lives" (Walberg 1977), "teacher beliefs" (Burn 1992; Richards 1998), beliefs, attitudes and knowledge (Woods 1996), "conceptions of teaching" (Freeman & Richards 1993), "conceptions of practice" (Freeman 1996), "teachers' maxims" (Richards 1998), "teachers' pedagogical systems" (Borg 1998), "teachers' theories" (Borg 1999) and

"teachers' personal theories" (James 2001). In a review of the literature, Borg (2003) identified sixteen such terms for teacher cognition that have been in use over the past decade. Despite the different nuances in the meanings of these terms, they do share a common premise: "a teacher's cognitive and other behaviors are guided by, and made sense of, in relation to a personally held system of beliefs, values, and principles" (Clark & Peterson 1986:287). This study therefore adopted an interpretative perspective on teachers' beliefs. From this perspective, teachers are seen as knowing, meaning-making beings, and hence knowledge and meaning-making have an influence on their actions.

2.1.2. Factors that inform teachers' beliefs

Borg (2003) provides a framework for understanding some of the major influences (Figure 1). Figure 1 shows that the main influences upon teacher cognition have been found to be schooling, contextual factors, professional coursework, and classroom practice, with the latter two being in a reciprocal relationship to teacher thinking; that is, a teacher's beliefs, knowledge, etc. both affect and are affected by the teacher's experiences in professional coursework and in the classroom. If teacher assessment, as one of the cognitive processes that teachers engage in, is connected to a wider network of beliefs, principles, etc., then it follows that teacher assessment is also influenced by a variety of factors. One major goal of this study is to understand what those factors are and how they impinge upon ESP assessment thinking.
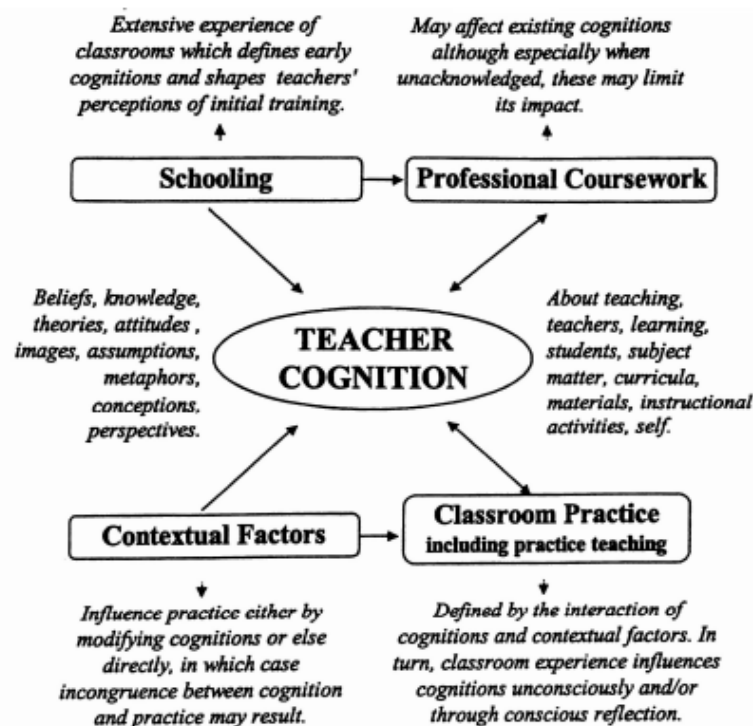

Fig. 1 Influences on teacher cognition (from Borg, 2003:82)

Beliefs may also be mediated by some contextual factors which have a great influence on the formulation, modulation, adjustment, and changes to the beliefs teachers might hold, and this affects their approach to putting theory into practice. Contextual factors can play one of two roles, either as facilitating agents for teachers to further their work, or as obstacles in their teaching-learning atmosphere, impeding the implementation of teachers' beliefs. Wood et al. (1991) contend that teachers' beliefs could be mediated by the differences in the nature of the content area and sometimes by the instructional materials available to them (cited in Kagan, 1992).

Some research has revealed the interconnection between the different beliefs of a teacher. Munby (1984) contends that teachers' beliefs come from their own experiences, and that instructional principles are pragmatic rather than theoretical. For example, sometimes teachers in their study did not articulate a well-grounded teaching theory. Instead, they upheld pragmatic views by saying that they wanted to "make learning fun" and that they employed "a different way of learning". This pragmatic view of teaching also influences their views of the curriculum and their evaluation of teaching. They evaluated their teaching according to how much their students had learned and remembered. This shows that different beliefs within an individual teacher are interconnected. Pajares (1992) reminded us that when studying teachers' beliefs, it is important to consider the connections among beliefs, rather than regarding them as independent systems. He further points out that research on teachers' beliefs should forge connections between different beliefs, beliefs and teacher practices, and beliefs and student outcomes.

Generally speaking, factors affecting teachers' beliefs can be subjective and inner, or objective and outward. One's own previous experiences are subjective and inner factors, and often contribute to one's tenacity and a priori conditions for perceiving things. On the other hand, opinions from authorities or 'important others' and what has been tested in scientific research are objective and outward factors affecting teachers' beliefs. Many of the studies examining teachers' principles have considered general matters guiding classroom conduct and management. Studies of teacher thinking in specific areas (e.g., reading) have also shown that teachers abide by certain principles (Borg 2003). It is thus reasonable to assume that principles also guide teachers' thinking in assessment. Remesal's (2011) qualitative analysis of fifty teacher interviews suggested that teachers' assessment conceptions (an organized system of beliefs) may in fact consist of different and even contrasting beliefs about the various roles of assessment (e.g., assessment's role in teaching versus its role in learning). Thus, it is important for researchers and teacher educators to consider distinct assessment beliefs and not assume, for instance, that beliefs about what to assess will align with beliefs about the tasks used to assess.

2.2. Classroom assessment

The way we see learning taking place is crucial to how we construe teaching as an activity, but it is also crucial to how we construe assessment. Obviously, assessment should not be an isolated activity operating independently of, and therefore without impact on, teaching. On the contrary, teaching, learning and assessment are inextricably interrelated.

2.1.1. Definitions of classroom assessment

Airasian (1991) defines classroom assessment as "all the processes used by teachers for collecting information and for making interpretations and decisions based on this

information on a daily basis in the classroom in order to improve teaching and learning" (cited in Mavrommatis 1997:381). In this sense, what teachers do in classroom assessment is make decisions in the hope of facilitating learning and teaching. Herman et al. (1992) pointed out that the decisions of teachers based on the information provided by classroom assessment include "what students have learned, what grades are deserved, whether students should be allowed to move on to the next grade, what group they should be assigned to, what help they need, what areas of classroom instruction need revamping, where the school curriculum needs bolstering, and so forth" (p.95). Stiggins (1991) stated that classroom assessment serves at least three sets of purposes. First, it serves as a tool for making informed decisions, including diagnosing students' needs, grouping students for instruction, and assigning grades, etc. Second, it serves as a teaching tool. For example, through classroom assessment teachers can communicate their expectations of a student's performance with that student, provide students with a chance to practice, and have them engage in self-or peer-evaluation so as to help them become better performers. Third, it serves as a tool for classroom management and for keeping students in line (Stiggins 1991). From these definitions and uses, we can see that classroom assessment is not simply assigning a grade, and that it permeates all stages of instruction (Brookhart 1997). The above confirms that classroom observation is a useful way to capture the complexity of the ESP practices of teachers.

2.2.2. Dimensions of classroom assessment

Educational assessments fall into two general categories, summative and formative assessment. Summative assessment was defined by Rolfe and McPherson (1995) as "mandatory, formal, and given at the end of a prescribed period of instruction" (p.837). They then went on to indicate that a summative assessment requires students to demonstrate the "sum" of their knowledge acquired over a period of time. Biggs (1998) however, suggests that a summative assessment should also be seen as formative. That is, it should be used as a learning tool, not simply viewed as the final part of the learning cycle for students.

The second category of educational assessment is formative assessment. A search of the literature revealed a plethora of terms and methods used to indicate the activity of measuring student learning that provides feedback to students during the learning cycle. Some of those terms include: classroom assessment (Brookhart 1997), embedded assessment (Gallagher 2000), formative evaluation (Fuchs 1995), and alternative assessment (Chen & Martin 2000). The defining difference in summative and formative assessments is that a formative assessment in its purest form is an assessment that occurs during learning, provides non-judgmental feedback to students, and helps narrow the gap between what a student knows and what the intended objectives of learning are (Black & William 1998a, 1998b). The above discussions indicate that the scope of a classroom observation should not be limited to summative assessment. Therefore assessment in this study refers to both summative and formative assessment which, following Stiggins (2002), can be seen as an assessment *of* learning and an assessment *for* learning.

In Knowing What Students Know: The Science and Design of Educational Assessment (National Research Council 2001), the authors describe a model for teachers' reasoning about assessment that is referred to as an ''assessment triangle.'' This triangle represents essential components of the design and implementation of an assessment by a teacher. The three essential elements are: (1) Cognition - the teacher's theories about what students know, how they learn, and what is important to understand in the subject domain, (2) Observation -

the teacher's assumptions about which tasks or situations will elicit and reveal students' understanding, and (3) Interpretation - the teacher's set of premises and interpretive methods for drawing inferences from collected evidence of student learning (see Figure 2).

Collectively, the essential elements of this model predict how teachers' decisions about assessment derive from a set of complex interacting factors that shape and are shaped by the learning environment in which the reasoning and decision-making occurs. From my perspective, the model can be a useful construct for inspecting teacher thinking and decision-making with regard to assessment. In particular, I will focus my attention in this work on the Cognition corner of the triangle and explore the different factors that ESP and subject teachers consider when making decisions about tasks designed to formatively assess student understanding. I am particularly interested in the characterization of the features that teachers use to make assessment decisions and the analysis of how these factors differ among ESP and subject teachers with different levels of preparation and experience.
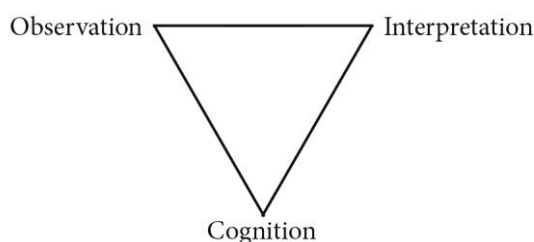
Observation ▽――――――▽ Interpretation

Cognition

Fig. 2 The assessment triangle
(National Research Council. 2001, Knowing what students know (p.44)
Washington, D.C.: National Academy Press)

2.3. ESP testing

Until recently, there has been very little in the way of research or publications for those who need to assess English for specific purposes. Teachers and testers have had to take what has been produced for teaching purposes, and seek to convert and adapt it for assessment. There has been very little practical guidance for test development, and there has been no attempt to develop a theoretical framework within which the assessment of English for specific purposes might develop (Dudley-Evans & St John 1998).

In a discussion of ESP testing, Douglas (2010: 33) pointed out that over the years, language specialists have made the following arguments about ESP testing: 1) Specific purpose language proficiency is really just general purpose language proficiency with technical vocabulary thrown in; 2) we don't need specific purpose tests since, if we test general language knowledge, specific uses will take care of themselves; 3) specific purpose language tests are unreliable and invalid since subject knowledge interferes with the measurement of language knowledge; 4) there is no theoretical justification for specific purpose language testing; and 5) specific purpose language testing is impossible anyway, since the logical end of specificity is a test for one person at one point in time. Douglas (2010) disagrees with these assertions, and argues that (1) language acquisition is a special case of a general capacity for language use, (2) both dialects and registers are learned and discarded as part of social behavior – this is related to the discourse domains hypothesis, and (3) are learned in contexts, so that the interaction between language knowledge and context changes the nature of both. Thus, it surely must be the case that there is such a thing as specific purpose language.

The following presents three views of ESP testing. Firstly, the field of ESP testing has been seen as a separate and distinctive part of a more general movement of English language testing, focusing on measuring specific uses of English language among identified groups of people, such as doctors, nurses, lawyers, civil engineers, tour guides, air traffic controllers, and others. Secondly, ESP testing has been viewed in the broader context of the teaching and learning process. From the perspective of Dudley-Evans and St. John (1998) assessment does not stand alone, but occupies a prominent place in the ESP process, giving an ESP teacher a wealth of information on the effectiveness and quality of learning and teaching. As shown in Figure 3, assessment interacts with needs analysis, and is dependent on course (and syllabus) design. Thirdly, tests enhance the learning process and act as a learning device. Put in the words of Dudley-Evans and St John (1998: 212), an ESP test is "an aid to learning". Moreover, assessment evaluates the benefits of learning, as tests can give learners a sense of accomplishment and a feeling that the teacher's evaluation matches what skills and knowledge have been covered. Along the same lines, Richards and Renandya (2002) suggest that assessment yields an observed judgment of the effectiveness of teaching. Dudley-Evans and St John (1998: 210) add that assessment "encompasses benefits such as reinforcement, confidence building, involvement and building on strengths".
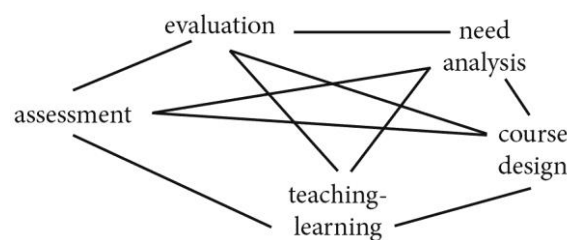

Fig. 3 Stages in the ESP process (Dudley-Evans and St John 1998: 121)

2.3.1. Assuring the quality of ESP tests

Douglas (2000) argues that authenticity is central to ESP testing. The underlying assumptions here are that 1) language use and performance can vary according to context, 2) specific contexts have distinguishing lexical, semantic, syntactic and phonological features, and 3) these features can only be realistically approached through the use of authentic test material.

Bachman (1990), writing in the context of language testing, develops this idea further with his distinction between **situational authenticity** and **interactional authenticity**. The situational aspect is composed of authentic characteristics derived from a needs analysis of tasks in the target language use situation, the features of which are realized as test task characteristics. Thus, situational authenticity can be demonstrated by making the relationship between the test task characteristics and the features of tasks in the target language use situation explicit. The interactional aspect of authenticity involves the interaction of the test taker's specific purpose language ability with the test task. The extent to which the test-taker is engaged in the task (by responding to the features of the target language use situation embodied in the test task characteristics), is a measure of interactional authenticity.

Drawing on Bachman's (1990) conception of authenticity, Douglas (2000) proposes to employ his dual notion of authenticity in specific purpose language testing. According to Douglas, in ESP test development, what we must do is first describe a target language use

situation in terms of features of context and task. We must then specify how these characteristics will be realized in the test so as to engage the test taker in test tasks, performance on which can be interpreted as evidence of communicative language ability with reference to the target situation.

Another important factor in ESP testing is *backwash* or the *washback effect*. The notion of backwash is that of impact, the relationship between test use and the ESP situation in which it is used (Hughes 1989). Put simply, it is the effect of testing on classroom instruction, on "what is taught and how it is taught" (Dudley-Evans & St John, 1998: 214). Tests, especially those that are important for the test takers, may generate positive or negative backwash. The research that has been conducted in this area (Karim 2002) indicates that the relationship between testing and teaching is very complex and it is misleading to claim that good tests will automatically have beneficial effects on classroom instruction. Other variables such as teacher competence, motivation and innovation, the climate of the school, and the socioeconomic status of pupils and teachers combine to exert an equally important influence on what goes on in an ESP classroom.

However, if a test is regarded as important, then preparation for it can dictate teaching and learning activities. More importantly, a public examination might be a signal to students (and teachers) of what is important and what is not. An ESP test, based directly on an analysis of the English language needs of a specific group of learners as similar as possible to what would be required in real life, would probably be a more plausible case of beneficial backwash than a case in which no such analysis was carried out.

In short, what matters in ESP testing is whether learners can communicate using the field specific target language and use knowledge of the field in order to achieve their aims, in order to understand and be understood, in order to get their message across in English.

2.3.2. Frequent problems in constructing ESP tests

One frequent problem is that many ESP test developers do not have a detailed knowledge of the specialized subject they are writing a test for. To compensate for their lack of background knowledge, the opinions of expert informants are often indispensable. Other ways of analyzing language use in a specific domain by using context-based research (Douglas & Selinker 1994), and grounded ethnography (Frankel & Beckman 1982) are also available. In both cases, considerable time is needed to do an adequate task of assessing how language is used in a specific target situation and translating that knowledge into test items.

Another problem in some ESP tests is that raters lack expertise in the specialized domain of knowledge being evaluated. Depending on the nature of the response examinees are expected to give and the testing environment, this can be a potentially serious threat to validity. Considering the entire test development process, Douglas adds, "…the most difficult aspect of producing test specifications is making the leap from the analysis of the target language use tasks to the specifications of test tasks" (2000:113).

In reviewing these studies, I have reached the following conclusions about teachers' knowledge, practice, and development in ESP assessment. First, the tension between teachers' implicit knowledge of assessment and the prescribed standards calls for a new methodology to address the uniqueness and individuality of teachers' assessment practice. Second, the salient issue of teachers' actual practices in assessment and the limited understanding of such practices require further empirical research. The interactive nature of teachers' practice also highlights the importance of situating teachers' assessment practice in larger socio-cultural contexts. Third, many of the ESP instructors are not domain experts, which makes the assessment task all

the more daunting. Based on these conclusions from the literature, this study explores the perceptions of teacher assessment through a qualitative inquiry into the experiences of six teachers who were involved in ESP courses.

3. METHODOLOGY

3.1. Theoretical underpinnings

An investigation of the beliefs held by the teachers themselves requires the research to be conducted from a "second order" perspective. The way teachers think about assessment cannot be observed. A second-order qualitative research approach known as phenomenography (Marton 1981, 1986) which has been used extensively in research on student learning, was adopted for this study. It is arguably best suited for exploring teacher beliefs of ESP assessment as they present and reveal themselves as dialogical phenomena in ESP instruction. It is capable of conducting that exploration by studying the meanings of the individual and the inter-subjective experiences of ESP teachers. It is a descriptive, qualitative, and reflective approach which analyzes the implicit meanings that characterize a phenomenon, and seeks the interrelationships between these meanings in the form of a meanings structure. Participants' descriptions of their experience, with their perceptions and understandings of how they underwent it, are essential to uncovering the nature of ESP assessment as implemented in ESP courses.

From the second-order, phenomenographic perspective, beliefs are not hypothesized to reside within individuals, but are relations between individuals and a particular task and context. Consequently, they are not stable entities within cognitive structures but are dynamic and depend on the particular context and task in which they are being studied. At the same time though, it is assumed that there are only a small number of conceptions and approaches about particular phenomena, and these can be identified and described (Marton 1981, 1986). Thus, this particular approach to research means that beliefs need to be identified and described within particular contexts, in terms of particular tasks and from the perspective of the teacher or learner within that context engaged in a particular task. This study has concentrated on the beliefs of ESP assessment held by six teachers at four Taiwanese universities.

3.2. The participants

The participants of this study are six teachers teaching ESP courses at four technological universities in Taiwan. According to Daiute and Fine (as cited in Josselson et al. 2003), the search for meaning implicitly depends on the collection of multiple perspectives. Sample sizes in qualitative studies differ depending on the purpose of the study and the specific qualitative methods used (Creswell 1998; Sandelowski 1995). The researcher followed the view that it is the quality of the data obtained rather than the quantity that is important (Sandelowski 1995). Because of the specific population the researcher wished to study, she did purposeful sampling (Merriam 1998) and criterion-based selection (LeCompte and Preissle 1993). Purposeful sampling is based on the assumption that the investigator wants to discover, understand, and gain insight, and therefore must select a sample from which the most can be learned (Creswell 1988: 61). In criterion-based selection you "create a list of the attributes essential" to your study and then "proceed to find or locate a unit matching the list" (LeCompte and Preissle 1993:70). The researcher was aware of the limitations of this study, and took advantage of rapport which had been established by her previous experiences with the participants. In order to focus on the

beliefs, practices, and instructional decisions of individual teachers, the author decided to select teachers representing varying characteristics such as sex (M or F), experiences (>2 years in profession), expertise (English teacher or subject teacher), and education (MS, MA, or Ph.D.) The selection was made since it was appropriate to her initial research questions regardless of the sampling issues. Table 1 presents relevant information about the participants. The researcher is aware that this sample cannot be regarded as representative of all ESP teachers in Taiwan; the findings may nevertheless be substantively applicable to other teachers in similar settings.

Table 1 Information about participants

Teacher	Gender	Years of teaching ESP	Education
ST 1	M	3	M.S. in Computer Science
ST 2	M	2	Ph.D. in Animal Science
ST 3	F	4	Ph.D. in Nursing
ET 1	F	3	Ph.D. in Media Culture
ET 2	M	5	MA in Technical Writing
ET 3	F	4	MA in TEFL

3.3. Data collection

Data for this research consisted of the following: 1) in-depth, informal, e-mail communications and oral interviews with the participants, individual interviews with ESP and subject teachers that were all tape recorded and later transcribed for coding and analysis; and 2) teachers' personal records, including their teaching and assessment plans, and the notes they made in relation to their assessment practice and their students' responses. The process adopted for collecting the data is as follows:

1. A semi-structured background interview with each of the teachers was conducted at the beginning of the fall semester in 2011. The purpose of this interview was to obtain basic background information about the participants and to understand their initial views on ESP teaching. An appointment for an interview was made with each participant, and an interview packet was given in advance. The packet consisted of a cover letter stating the purpose of the study, an informed consent form, and a list of interview questions.
2. A semi-structured follow-up interview with each of the teachers was conducted at the end of the fall semester using the event-recall technique, prompted by their own classroom assessment material. The interviews were designed to elicit reflections and descriptions of how the teachers carried out their ESP assessments with a view to exploring their interactive decision-making in their improvisational teaching performance with specific reference to ESP classroom assessments.

3.4. Data analysis

Interviews were transcribed utilizing the methods, procedures, and practices of phenomenological analysis by Moustakas (1994), the modified Van Kaam method. The specific steps utilized were as follows:

1. Transcriptions of the interviews
2. Coding the expressions relevant to the experiences was conducted.
3. Patterns and themes were identified.
4. The uncovering of the meaning or meanings of the statements was conducted.

5. From the uncovering of the meanings, common categories, patterns, and themes were deciphered.
6. A comprehensive thematic description of the experiences was developed from the common categories, patterns, and themes.
7. A textural-structural synthesis of the meaning and essence of the experiences of teachers' ESP assessment practices in Taiwan was constructed.

4. FINDINGS

In general, the results showed that there was considerable variation in the teachers' descriptions concerning the purposes of ESP assessment and the practices they apply. The categorization of the teachers' descriptions of ESP assessment appears below in two parts. The first part addresses the teachers' conceptions of ESP assessment, while the second part is about the tendencies in ESP assessment practices mentioned during the interviews.

4.1. Beliefs about ESP assessment

In order to reveal the participants' beliefs regarding ESP assessment, the ESP and subject teachers were asked to use metaphor to describe what ESP assessment is and to give a reason for their choice. The metaphors are important for learning teachers' beliefs of assessment because as Lakoff and Johnson (1980) pointed out "our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature" (p.454)

Even though there were 6 participants, 11 different metaphors were produced in response to the question "What is the image or metaphor of ESP assessment?" The participants gave more than one metaphor. The varied responses that represent these teachers' beliefs about ESP assessment can be classified into three main categories: assessment as *inquiry*, a *summative process*, and *something agitating*.

4.1.1. Assessment as inquiry

Five of the 11 metaphors fell into the assessment as inquiry category. These show that assessment is seen as an ongoing process (*lifelong learning, deep-sea diving*), as diagnosing learners' weakness and strengths from different points of views (*a window with a lot of angles*), and as a way of promoting learning through feedback (*a road with traffic lights, a close friend who shows you the way*), all of which are indicators of formative assessment practices (Black and William 2004; Tunstall and Gipps 1996). These teachers do not limit their assessment practices to observing students' linguistic proficiency; they use assessment as an integral part of the learning process. As seen in Table 2, both groups of participants produced a similar number of metaphors for this theme. It is possible to assume that, regardless of academic background, teachers are aware of the positive effects of formative assessment practices on learners' progress.

4.1.2. Assessment as a summative process

The metaphors that were grouped under the category of *assessment as a summative process* were indicators of summative assessment: assessment that shows the end product *(a story because it ends either well or badly for the students)* and assessment to grade learners' products *(like a mirror because it shows us how much students learnt, like a scoreboard*

because it shows students' situation regarding course achievement.). It is not surprising to encounter such themes, as teacher need to have an overall view of their student performance, which indicates summative assessment. Moreover, in an exam-oriented country such as Taiwan, where this study took place, emphasis is placed more on the scores as outcomes and measures of abilities. However, analyzing metaphors separately for each group shows that academic background might have a minor role on teachers' assessment beliefs (Table 2). Assessment was least likely to be conceived as a summative tool by the ESP teachers, as only 1 out of 6 of their metaphors was related to this.

This may be due to the fact that the subject teachers' were exposed to effects of the traditional teaching and learning processes when they were students, which, in turn, may have formed their beliefs that learning is a process which requires students to use cognitive processes only. Although this explanation is reasonable because teachers' teaching and learning-related behaviors and beliefs are significantly affected by their earlier educational experiences when they were students (Pajares 1992; Rodgers and Scott 2008), it requires further investigation due to the reason that the subject teachers' conceptions of teaching and learning were not examined in this study. Indeed, if subject teachers perceive ESP assessment as a summative process through which students' learning is evaluated as a product, or as an end-state, of their class-related tasks and activities, then it is not surprising to observe that they also perceive the assessment process as a competitive process because summative evaluation required teachers to evaluate students' learning performances on the basis of highly structured and school-based extrinsic norms. It is obvious that such assessment, referred to as a "one size fits all" approach, provides a basis for students to compare their performances with each other rather than to focus on their performances as the degree to which they have made progress in their learning.

4.1.3. Assessment as something agitating

Conceiving *assessment as something agitating* was the third most frequent metaphor that emerged from the data. Metaphors associated to this were mainly about fear (*a cloud because you never know what it brings*) and difficulty (*a road of struggle; a difficult math problem*). To understand this negative connotation of assessment, it might be necessary to consider how 'testing' is generally conceived. In a conference, Alderson (2005) defined language testing as something technical, secret, full of forbidding jargon, remote from teachers and teaching, and thus ignorant of the goals and the realities of the language classroom. It is very common to have such negative beliefs about testing; however, it is necessary for teachers and learners to make a distinction between the terms of testing and assessment. Unlike testing, assessment, as Alderson puts forth, helps teachers establish where their learners are at present and what level they have achieved, give them feedback on their learning, and diagnose their needs for further development. Though not very high in percentage, considering assessment as something agitating probably occurred because of confusion between testing and assessment.

The existence of negative metaphors regarding the assessment concept is regrettably meaningful and should be given serious attention during teacher education because evaluation is one of the crucial aspects of education (Nitko 1996). Thus, it can be said that teacher educators should examine why teachers have negative metaphors about the assessment concept, such as clouds, a road of struggle, and a difficult math problem, in order to turn these negative thoughts into positive ones during teacher education.

Table 2 Metaphors of teachers' beliefs about ESP assessment

	ESP Teacher (number of participants)	Subject Teacher (number of participants)
Inquiry	3/3	2/3
A summative process	1/3	2/3
Something agitating	2/3	1/3

4.2. Reported ESP assessment practices

4.2.1. Functions of assessment

Collectively, the interviewees described three basic functions of assessment in their ESP courses:

    1) initial assessment, prior to courses beginning;

    2) ongoing, formative assessment in relation to writing tasks; and

    3) assessing students' achievements during or upon completion of a course.

The author was surprised to find that these assessment functions took a somewhat different realization for courses taught by ESP or subject teachers, as summarized in Table 3 and as the following examples from the interviews indicate.

Functions of initial assessment

Initial assessment was seldom mentioned by the subject teachers. Seemingly, they assumed that students' enrolment in a particular academic field or type of program defined *a priori* their needs for the ESP course, so decisions about needs or diagnosis were made generally on the basis of program policies. In contrast, ESP teachers described procedures for initially assessing students' English abilities for the purposes of diagnosis to inform teaching or course design.

The students we taught are diverse level proficiency of adult learners. We need to know their level before we design our course. (ESP teacher 1)

Functions of formative assessment

Most accounts of assessment during the interviews focused on ongoing, formative assessment of students' English and the grading of it. Subject teachers tended to emphasize the realism and value of the tasks selected for writing as well as the importance of assessment adhering to standards, for example:

The mid-term is worth 30%, the final exam is 30%, and their class work is 40%. Their written papers for these follow the same materials as the material they study and practice writing: to fill in forms, to make job applications, to write resumes, do interviews, which they need practice in and techniques for , and they take notes from phone conversations. (Subject Teacher 1)

Are the students able to express themselves, correctly? Certain key words, did they use them? ... So my challenge is to make sure that they ... understand... and that they are able to express their ideas. (Subject Teacher 3)

Similarly, many instructors described intricate ways of integrating assessment with ESP curricula. This concern was expressed particularly by the ESP teachers, where the instructors focused intently on the individual development of students (compared to the subject teachers):

Assessment is consciously integrated with the syllabus. The purpose of assessment is to make sure teachers are assessing what they are teaching and the students are learning. There are benchmark texts and tasks with minimum standards that students should attain. Teachers often give students the performance criteria so students can monitor their own progress. Teachers design assessment tasks that reflect what is taught and what learners need to learn. The assessment should fit into the learning context. They shouldn't notice the bump when they get to assessment. (ESP Teacher 2)

The student's individual progression must be taken as a basic reference criterion for accrediting and in order to undertake changes in teaching. (ESP Teacher 1)

This personalized focus on individual students seemed to prompt instructors to use formative assessment as a basis for record-keeping (with reference to individual students) and instructional planning (with reference to groups of students), particularly for ESP teachers. The author was surprised to observe that this tendency was seldom cited in the discussions of the subject teachers, where results from prior needs analyses or program policies appeared to define curriculum content and initial entry requirements in linguistic or rhetorical, rather than personal, terms.

I keep marks to remind myself about the individual students, for example, if they are using appropriate vocabulary, grammar, organization, if they got the message across. I keep this in a book for myself in planning my teaching and to see if they make progress along the way. (ESP teacher 1)

For feedback on their assignments, I respond to their ideas, not correct them. There are some common problems among the students, for example, no grasp of paragraph structure, they don't know how to support their ideas, they copy a lot, and have problems with certain lexical items. I discuss these and give examples of them in class. (ESP teacher 3)

Table 3 Tendencies in assessment practices mentioned during interviews:
ESP teachers vs. subject teachers

	ESP Teacher (number of participants)	Subject Teacher (number of participants)
Functions of initial assessment		
Diagnosis for instruction	(3/3)	(1/3)
Functions of formative assessment		
Helping students improve	(3/3)	(3/3)
Maintaining standards	(0/3)	(3/3)
Linking assessment into curriculum	(3/3)	(3/3)
Record-keeping	(2/3)	(0/3)
Methods of assessing achievement		
Tasks based on prior needs analyses	(2/3)	(3/3)
Proficiency tests	(2/3)	(1/3)
Rating scales and competency tests	(1/3)	(3/3)
Grading tasks	(0/3)	(2/3)
Combinations of assessment methods	(3/3)	(1/3)
Achievements observed among students		
Language and style	(2/3)	(0/3)
Self-confidence and expressive abilities	(3/3)	(1/3)

4.2.2. Methods of assessing achievement

For the pedagogical function of assessing students' achievement, distinctive differences appeared between ESP teachers and subject teachers. The subject teachers defined their standards for achievement in their own terms, deriving them from prior needs analyses and the constructs guiding the syllabi:

We do competency-based assessment. Analyses of register and appropriateness define the task and give the criteria. If the student achieves the task, then they are certified for having done it. If it is not achieved or only partly achieved, they resubmit it. (Subject Teacher 3)

They do a simple project, first design it, second create a questionnaire, third collect data and analyze it, fourth make an oral presentation, then produce a written report. At the end of the course they present the final report, and it is marked when completed, which forces them to see it as a whole. I provide a list of criteria that are assigned holistically, not discretely. (Subject Teacher 2)

Grading tasks was yet another method of assessing achievement commonly described by subject teachers:

I do my grading based on how successfully they have done the task, convinced me, analyzed and evaluated and presented the material in an interesting and unique fashion, and how they have put themselves into it, their own writing style, amplified that style, worked on coherence, and especially vocabulary. I try to see how students can identify how to expand their vocabulary to provide more expressive and unique vocabulary. (Subject teacher 1)

Continuous assessment is the typical practice – no exams. The feedback to students is extensive on their papers, and students seem to heed it, and the grade follows. (Subject teacher 3)

In contrast, a variety of different methods were described for assessing achievement by the ESP teachers:

Combinations of assessment methods, moreover, were cited by two of the ESP instructors as ways of determining student achievement in the ESP courses:

I take students through a series of tasks that contribute to their final mark. Some are done in class, so I can check, if in doubt. Then there is a comprehensive test at the end, aiming to reinforce all the skills taught, for example, recognizing order of presentation, conclusion, referencing. For the 2 main essays there is a grid of criteria, for example, how well they have answered the essay question, how relevant the material is, whether they have used evidence, the development of an argument, and their presentation, including language. (ESP teacher 2)

Fifty-five percent of their grades are for their portfolios, which include their major assignments, media essay, and a letter about their personal learning process, plus 5 to 10 other pieces they choose, which may be an assignment for another course. Twenty percent is for the final exam, which includes 2 essays of 300 to 500 words, marked for evidence of planning, organization, writing skill, and editing. I would prefer not to have an exam, but this is the common way of doing things at this university. (ESP teacher 3)

4.2.3. Achievements observed among student

Nonetheless, the author was surprised at the distinctions in ESP and subject instructors' conceptualizations of student achievement overall. When asked what achievements they saw their students making in their ESP courses, subject instructors focused on a relatively narrow range of observable behaviors. As noted above, these outcomes were defined:

They go out being able to write an excellent letter of application. They understand the concept of brevity, not plagiarizing, how to do referencing, bibliographies, footnotes for a report, and they have their CVs on computer disc and keep good, fault-free versions. (Subject teacher 2)

They make achievements in the organizational aspects of their writing. They get the idea that information has to be structured to be presented effectively. But 3 months is not enough to learn much language. (Subject teacher 1)

In contrast, when the ESP teachers were asked the same question, they mentioned a wide range of differing indicators of achievement. These included improved language and style, increased self-confidence and expressive abilities.

a. *Language and style* were areas of achievement cited by two ESP instructors:
   It is rare in an 18-week course to see quantum leaps, but I am still surprised by the predictable pathways in development, for example, many students are able to write more coherent, satisfying conclusions, are more aware of morphemes and tense and appropriate use of linking words, are better at controlling sentence length and complexity, better paragraphs, referencing ideas. There are so many small things that go together to make good writing. (ESP teacher 1)
b. *Improved self-confidence and expressive abilities* were cited by three ESP instructors as indicators of achievement in ESP courses:
   At the advanced level, they show an expanded range of writing and are able to express themselves forcefully and coherently. They seem to be more unique in the way they write. They move away from a standard way of writing, for example, having to do an introduction and conclusion, and realize that the nature of the task defines this; it is not a pat formula. (ESP teacher 3)

5. CONCLUSIONS

Analysis of the metaphors produced by the teachers in this study revealed that the participating teachers view assessment as being embedded within instruction. In other words, they perceive assessment as a way to provide evidence of teaching and learning. However, analyzing metaphors separately for each group shows that academic background might have a minor role in teachers' assessment beliefs (Table 2). As Rea-Dickens (2004) states, teachers are agents of the assessment and this role is important in the washback effect of assessment (p.252). Reported practices revealed ESP and subject teachers' beliefs about ESP assessment. The data from this study support the result of previous studies (e.g., Brown et al. 2011; Remesal 2011; Yin 2010). The orientations that teachers adopt to conceptualize curricula for ESP instruction appear to influence the ways in which they assess their students' achievement. This distinction was especially prominent with respect to the methods of assessment the teachers said they utilized and the types of achievements in ESP courses they perceived their students as making.

The subject teachers took a Hymesian view of communicative competence by considering language proficiency in terms of real world criteria (i.e. are teachers creating the necessary conditions for classroom learning to take place?) which, as subject teachers, they felt well-qualified to assess. The capabilities that the subject teachers assessed requires that students share some of the capabilities of the real-world practices the course is meant to prepare them for – but it is not the same as those practices. In implementing them they come closer to what McNamara (1990) defines as the 'strong' approach to performance testing whereby language is assessed in terms of successful task completion, with all that it entails. English teachers on the other hand veer towards the 'weaker' (and arguably more conservative) approach by focusing more closely on what they are trained to assess, that is, the quality of the language sample elicited through the teaching task.

The implication is that subject teachers might be more concerned with the 'correctness' or 'worthiness' of the subject content. This was in fact evident in a comment made by one of the subject teachers: "my judgments of his language ability are clouded by the way he presents the topic". It is quite conceivable that in assessing use of subject specific language, the English teachers are focusing on the lexis, grammar and internal cohesion of the presentation while the subject teachers are more concerned with the way in which the subject content is conceptualized. While it is generally accepted that subject teachers should be consulted during the needs analysis phase of specific purpose language test development, their role in the actual assessment process is seldom considered. The role of the subject teachers is a very significant one though and it should not be neglected in ESP classrooms. A degree of co-operation between the ESP teacher who implements the course and the subject teacher who acts as a monitor and advisor of the ESP assessment should be considered more appropriate.

The data which represents a group of teachers' beliefs and practices about ESP assessment in Taiwan has portrayed a partial picture for future investigation, though it may not be complete. The limitations of the present study go beyond its contextual and institutional constraints. Given the particularly complex and multi-faceted nature of classroom assessment and the myriad factors that can shape teachers' beliefs, no attempt will be made to generalize the findings of this study beyond the local context, although comparison data suggest feasibility of finding general trends across contexts, and individual differences of teachers, through replicated studies. A further concern is "the potential limitations of self- reported data" (Shavelson, Webb, and Burnstein, 1986: 44). From a phenomenological perspective, the teachers' language provided the best view of the meanings they attached to their experiences. However, the study was dependent on their willingness and abilities to describe their realities and reveal their "true" selves. Finally, there was no student voice in this study which could have allowed for important and interesting insight into the washback of ESP assessment on teaching and learning. However, due to the time constraints and the researcher's attempt to keep the data set manageable, the data only represented the teacher's perspective and thus limited the scope of the study.

While ESP assessment is a very complex phenomenon and there are many ways in which it can be approached, the findings of the current study have provided a number of possible directions for further research in this area. Future research into ESP assessment, by taking learner perspectives into account, will provide more grounded accounts of ESP assessment and its implications for test validity. In addition, to get a more comprehensive picture of ESP assessment, it is desirable to conduct studies which look at ESP assessment from different perspectives using different research methods (including the two central

participants involved: teachers and students) in order to investigate the influence it exerts on ESP classroom learning and teaching in depth.

REFERENCES

Airasian, Peter W. 1991. *Classroom assessment*. New York: McGraw-Hill.

Alderson, J. C., Clapham, C., and Wall, D. 1995. *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.

Alderson, J. C. 2005. *Principles and practice in language testing*. Keynote address at RATE-QUEST Conference, Cluj-Napoca, Romania, October 28-30.

Bachman, L. F., and Palmer, A. S. 1996. *Language testing in practice*. Oxford: Oxford University Press.

Bachman, L. 1990. *Fundamental considerations in language testing.* Oxford: Oxford University Press

Belcher, D. 2006. "English for specific purposes: Teaching to perceived needs and imagined futures in worlds of work, study, and everyday life". *TESOL Quarterly* 40(1): 133-156.

Binnie Smith, D. 1996. "Teacher decision making in the adult ESL classroom." In *Teacher learning in language teaching*, edited by Freeman, D. and Richards, J. C., 197 – 216, Cambridge: Cambridge University Press.

Biggs, J. 1998. "Assessment and Classroom Learning: a role for summative assessment?" *Assessment in Education* 5(1): 103-110.

Black, P. and Wiliams D. 2004. "Inside the black box: raising standards through classroom assessment." In *Assessment for Learning: An Action Guide for School Leaders,* edited by S. Chappuis, R. J. Stiggins, J. Arter, and J. Chappuis, 10-21, Portland, OR: Assessment Training Institute.

Black, P. and William, D. 1998. *Inside the Black Box: Raising Standards Through Classroom Assessment*. London: Department of Education & Professional Studies, King's College London.

Black, P. and William, D. 1998. "Assessment in classroom learning." *Assessment in Education* 5(1): 7-71.

Borg, S. 2003. "Review article: teacher cognition in language teaching: a review of research on what language teachers think, know, believe, do." *Language Teaching* 36: 81 – 109.

Borg, S. 1999. "Teachers' theories in grammar teaching." *ELT Journal* 53 (3): 157 – 167.

Borg, S. 1998. "Teachers' pedagogical systems and grammar teaching." *TESOL Quarterly* 32 (1): 9 – 38.

Brookhart, S.M. 1997. "A theoretical framework for the role of classroom assessment in motivation student effort and achievement." *Applied Measurement in Education* 10:161-180.

Brown, G. T. L., Hui, S. K. F., Yu, F. W. M., and Kennedy, K. J. 2001. "Teachers` conceptions of assessment in Chinese context: A tripartite model of accountability, improvement, and irrelevance." *International Journal of Educational Research* 50: 307- 320.

Burns, A. 1992. "Teacher beliefs and their influence on classroom practice." *Prospect* 7 (3): 56 – 66.

Carless, D. 2005. "Prospects for the implementation of assessment for learning." *Assessment in Education* 12: 39–54.

Chang, C. 2005. "Oral Language Assessment: Teachers' Practices and Beliefs in Taiwan Collegiate EFL Classrooms with Special Reference to Nightingale University." Ed. D. dissertation, University of Exeter.

Chen, Y. F., and Martin, M. A. 2000. "Using Performance Assessment and Portfolio Assessment Together in the Elementary Classroom." *Reading Improvement* 37(1):32-38.

Clark, C. M. and Peterson, P. L. 1986. "Teachers' thought processes." In *Handbook of research on teaching,* edited by Wittrock, M. C. 255 – 296, New York: Simon and Schuster Macmillan.

Creswell, J. W. 1998. *Qualitative inquiry and research design: Choosing among five traditions*. SAGE Publications, Inc. Thousand Oaks, CA.

Davison, C. 2004. "The contradictory culture of teacher-based assessment: ESL teacher assessment practice in Australia and Hong Kong secondary schools." *Language Testing* 21: 305–334.

Douglas, D. 2010. *Understanding Language Testing*. London: Hodder-Arnold.

Douglas, D. 2000. *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.

Douglas, D. and Selinker, L. 1994. "Research Methodology in Context-based Second-Language Research." In *Methodologies for Eliciting and Analyzing Language in Context,* edited by E. Tarone, S. Gass, and A. Cohen, 119-132, Hillsdale, NJ: Erlbaum.

Dudley-Evans, T. and St. John, M. J. 1998. *Developments in English for Specific Purposes*. Cambridge: Cambridge University Press.

Frankel, R. and Beckman, H. 1982. "IMPACT: an interaction-based method for preserving and analyzing clinical transactions." In *Explorations in provider and patient interactions*, edited by Pettigrew, L. Louisville, Kentucky: Humana.

Freeman, D. 2002. "The hidden side of the work: teacher knowledge and learning to teach – review article." *Language Teaching* 35: 1 – 13.

Freeman, D. 1996. "Renaming experience/reconstructing practice: Developing new understandings of teaching." In *Teacher learning in language teaching,* edited by Freeman, D. and Richards, J. C., 221 – 241. Cambridge: Cambridge University Press.

Freeman, D. and J. C. Richards. 1993. "Conceptions of teaching and the education of second language teachers." *TESOL Quarterly* 27(2): 193 – 216.

Fuchs, L.S. 1995. "Connecting performance assessment to instruction: A comparison of behavioral assessment, mastery learning, curriculum-based measurement, and performance assessment." *ERIC Digest, ED381984*. Retrieved October 30, 2010 from: http://www.ed.gov/databases/ERIC_Digests/ed381984.html

Gallagher, M. 2000. "The challenges facing higher education research training." In *Quality in Postgraduate Research: Making ends meet. Proceedings of the 2000 Quality in Postgraduate Research Conference,* edited by M. Kiley, and G. Mullins. Adelaide: The Advisory Centre for University Education, The University of Adelaide.

Giorgi, A. 1985. "Sketch of a psychological phenomenological method." In *Phenomenology and psychological research,* edited by A. Giorgi, 8-22. Pittsburgh, PA: Duquesne University Press.

Herman, J., Aschbacher, P.R. and Winters, L. 1992. *A Practical Guide to Alternative Assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Hughes, A. 1989. *Testing for Language Teachers.* Cambridge: Cambridge University Press.

Hutchinson, T. and Waters, A. 1987. *English for specific purposes: A learning centred approach.* Cambridge: Cambridge University Press.

James, P. 2001. *Teachers in action.* Cambridge: Cambridge University Press.

Josselson, R., Lieblich, A. and McAdams, D. P. (Eds.) 2003. *Up close and personal: The teaching and learning of narrative research.* Washington, DC: American Psychological Association.

Karim, M. A. 2002. "High Stakes Examination: Washback in the Language Classroom." Paper presented at the 28th conference of the International Association for Educational Assessment Reforming Educational Assessment to Meet Changing Needs. Hong Kong, China, Sept. 02.

Kagan, D. M. 1992. "Implications of research on teacher beliefs". *Educational sychologist* 27 (1): 65-90.

Lakoff, G. and Johnson M. 1980. *Metaphors we live by.* Chicago: The University of Chicago Press.

LeCompte, M.D. and Preissle, J. 1993. *Ethnography and qualitative design in educational research* (2nd ed.) New York: Academic Press.

Marton, F. 1981. "Phenomenography -Describing conceptions of the world around us." *Instructional Science* 10(2):177-200.

Marton, F. 1986. "Phenomenography: a research approach to investigating different understandings of reality." *Journal of Thought* 21(3): 28 - 49.

Mavrommatis, Y. 1997. "Understanding assessment in the classroom: Phases of the assessment process --The assessment episode." *Assessment in Education* 4(3): 381-399.

McNamara, T. F. 1990. "Item response theory and the validation of an ESP test for health professionals." *Language Testing* 7: 52-75

Merriam, S. B. 1988. *Case study research in education: A qualitative approach.* San Francisco: Jossey-Bass.

Moustakas, C. 1994. *Phenomenological research methods.* Thousand Oaks, CA:Sage

Munby, H. 1984. "A qualitative approach to the study of a teachers' beliefs." *Journal of Research in Science Technology* 21: 27-38.

National Research Council (NRC). 2001. *Knowing what students know: The science and design of educational assessment.* Washington DC: National Academy Press.

Nitko, A. J. 1996. *Educational assessment of students.* Englewood Cliffs, New Jersey, NJ: Prentice-Hall.

Pajares, M. F. 1992. "Teacher's beliefs and educational research: Cleaning up a messy construct." *Review of Educational Research*, 62(3):307-332.

Rea-Dickins, P. 2004. "Understanding teachers as agents of assessment." *Language Testing* 21: 249-258.

Remesal, A. 2011. "Primary and secondary teachers' conceptions of assessment: A qualitative study." *Teaching and Teacher Education* 27(2):472-482.

Richards, J. C. 1998. *Beyond training: Perspectives on language teacher education.* Cambridge: Cambridge University Press.

Richards, J.C. and Renandya, W.A. (eds.) 2002. *Methodology in Language Teaching: An Anthology of Current Practice.* Cambridge: Cambridge University Press.

Rolfe, I. and McPherson, J. 1995. "Formative Assessment: How am I Doing?" *Lancet* 345:8953.

Sandelowski, M. 1995. "Focus on qualitative methods: Sample size in qualitative research." *Research in Nursing & Health 18*: 179-183.

Shavelson, R., Webb, N. and Burnstein, L. 1986. The measurement of teaching. In *Handbook on teaching* (3rd Ed.) edited by M. Wittrock, 1-36. New York: Macmillan.

Stiggins, R. 2002. "Assessment crisis: the absence of assessment for learning." *Phi Delta Kappan* 83(10):758-765.

Stiggins, R.J. 1991. "Relevant Classroom Assessment Training for Teachers*." Educational Measurement: Issues and Practice* 10(1):7-12.

Strevens, Peter. 1988. "ESP after twenty years: A-Re –appraisal." In *ESP State-of-the-Art,* edited by M. Tickoo, 1-13. Singapore: SEAMEO Regional Language Centre.

Tsou,W. 2009. "Needs-based curriculum development: A case study of NCKU's ESP Program" *Taiwan International ESP Journal* 1(1): 77-95

Tunstall, P. and Gipps, C. 1996. "Teacher feedback to young children in formative assessment: a typology." *British Educational Research Journal* 22:389-404.

Walberg, H. 1977. "Decision and perception: New constructs for research on teaching effects." *Cambridge Journal of Education* 7 (1): 12 – 20.

Woods, D. 1996. *Teacher Cognition* in *Language. Teaching: Beliefs, Decision-making, and Classroom Practice,* Cambridge University Press.

Wood, T., Cobb, P. and Yackel, E. 1991. "Change in teaching mathematics." *American Educational Research Journal* 28(3): 587–616.

Yin, M. 2010. "Understanding classroom language assessment through teacher thinking research." *Language Assessment Quarterly 7*(2):175-194

Yin, M. 2005. "A progressively focused qualitative study of teacher thinking in English for academic purposes (EAP) classroom language assessment*."* Ph.D. dissertation, University of Bristol.