

Elementary school EFL teachers' beliefs and practices of multiple assessments

Yu-Ching Chan

National Taipei University of Education

ABSTRACT

English language teaching in Taiwan was formally extended to the elementary schools commencing in 2001, as a result of changes in the national English education policy. To effectively assess student learning in the revised curricular context, English teachers were required to use multiple assessments. This quantitative study investigated teachers' beliefs and practices of multiple assessments and explored the difficulties affecting their practices. A questionnaire comprising a self-report Likert scale, multiple choice and open-ended questions was completed by 520 elementary school EFL teachers from Northern Taiwan.

KEYWORDS: *Multiple assessments, EFL teachers' beliefs and practices, teaching and assessment, beliefs and practices of multiple assessments*

Introduction

In recent decades, assessment has gained increased attention in education. English language teaching at the elementary school level in many countries has resulted in vigorous discussions on assessment of EFL students' English language learning. In both local and global contexts, researchers and practitioners have sought and trialed appropriate and efficient assessment methods for evaluating and monitoring young EFL learners' progress in language learning (Chen 2003; Chern, Ruan, & Yeh, 2001; Gattullo, 2000; Hasselgren, 2000; Hsu, 2000, 2003; Johnstone, 2000).

Assessment refers to any method, strategy, or tool a teacher may use to collect evidence about student progress toward achievement of established goals. It is a process of collecting information and gathering evidence about what students have learned (Chen, 2003; Wishon et al., 1998). Assessment that is incorporated into the school curriculum design consolidates the goals of education and enables practitioners to reflect on the actual learning situation (Zahork, 1995), thereby benefiting students and requiring teachers to focus on working toward student progress (Wishon et al., 1998). In particular, the purposes and functions of assessment may be classified as follows: (1) understanding the strengths and weaknesses of students' learning ability, (2) assisting teachers in monitoring student learning progress, (3) evaluating students' learning, and (4) placing students in learning groups based on given institutional standards (Heaton, 1990; Popham, 1995). Assessment may be likened to a double-edged sword,

producing both positive and negative effects (e.g., washback) (Brown, 2001; Brown & Hudson, 1998; Chen, 2003; Hsu, 2003; Richards, 2002).

No single assessment is able to thoroughly cover the learning progress or achievement of students. The notion of "multiple" in multiple assessment can be interpreted from different administrative perspectives (Yeh, 2001): (a) formative assessment is administered during the learning process with the aim of using the results to improve instruction; (b) summative assessment is performed at the end of a course often for purposes of providing aggregate information on program outcomes to educational authorities (Brindley, 2001); (c) traditional test-based assessment involving written or oral tests; and (d) task-based assessment which are activity-based, and includes teacher-student interaction, student-student interaction, body movement, teamwork, portfolios, and other language requirement and activities.

In terms of content, assessment can be multiple, comprising vocabulary, sentence patterns, songs and rhymes as well as the skills of listening, speaking, reading, and writing. Assessment can also be multiple in terms of tools. It can involve the traditional paper-and-pencil test or use of multiple media such as computer, tape recorder, or video recorder. In terms of the tester, assessment can be multiple, engaging teachers, peers and self (Brown, 2001). Moreover, in test-based assessment, discrete-point tests which focus on linguistic forms and structures and integrative tests which emphasize comprehension and application of language can be used based on the content and purpose of the assessment.

To strengthen the developmental and individual appropriateness of assessment for young EFL learners, scholars and educators have raised the use of a multi-faceted framework of assessment that addresses children's uniqueness and diversified strengths (Neuman, Copple, & Bredekamp, 2000; Shohamy, 2000). Generally, the use of multiple instruments in language assessment is based on the assertion that "language ability is a broad and complex construct that cannot be fully measured by tests" (Shohamy, 1994, p. 135). Brady (1997) supports this assertion by declaring that different forms of assessment are better suited for different aspects of learning. This assertion is also supported by Neuman et al. (2000). Overall, they suggest the promotion of appropriate assessment strategies to boost children's learning and development and to overcome difficulties in obtaining valid and reliable indices of younger children's development and learning using one-time test administrations.

Assessment in Taiwan

In Taiwan, where the assessment of English language learning has been criticized for its strong emphasis on testing, which has led to a negative washback effect on language learning (Hsu, 2003), the educational reform movement has promoted the establishment of the Grades 1-9 Curriculum Guidelines, with specific principles on administering "good" assessment (Cai, 1998; Chang, 1996; Huang, 2001; Jiang, 2000; Ren, 2001; Wishon et al., 1998; You, 2000). Education reformers are gaining progress toward the achievement of a positive washback effect, whereby assessment procedures in a curriculum correspond to course goals

and learning objectives (Brown & Hudson, 1998).

In the past decades, school entrance exams which comprised reading and writing exerted a strong influence on teaching and learning in Taiwan, resulting in classroom instruction which focus only on these two skills. Consequently, without instruction aimed at developing speaking and listening skills, most students in Taiwan study English for years, but are not able to learn spoken English or express themselves fluently in English (Chen, 2003). Such classroom focus has resulted in a "teach to test" approach, preparing students for assessment rather than developing their skills within well-planned a curriculum based on student and institutional needs or demands. .

Corollary to this instructional challenge are administrative concerns on assessment dealing with the development of updated assessment programs that can be used for large-scale testing and dissemination of assessment results, mainly providing substantial relevant information to both students and stakeholders.

Thus far, the foregoing assessment challenges in Taiwan have been partially addressed through the development of appropriate multiple assessment procedures, which have constituted the Nine-Year English Curriculum Guidelines in Taiwan, apart from traditional assessment (Chen, 2003; Yeh, 2001). The guidelines which stipulate that teachers should use multiple assessment measures to monitor student learning aim at generating a positive washback effect, with the collaborative effort among scholars, educators, and the authority.

Most of the studies on assessment in Taiwan have focused on traditional testing (Chang, 1984). However, with the assessment reform promoted by educators and scholars in recent years, many teachers have become aware of multiple assessments, paving the way for the adoption of multiple assessments in mainstream in education. Task-based assessment was effectively demonstrated (Chen & Ruan, 2001); portfolio assessment was successfully implemented in an EFL elementary classroom, yielding effective development in students' writing skills and revealing the interface among learning, instruction, and assessment (Hsieh et al., 2000).

In 2001 Taiwan's Ministry of Education introduced the Guidelines for Grades 1-9 Integrated Curriculum which formally extended English language teaching to the elementary school level, requiring efficient assessment of language skills development through multiple assessment, beyond the traditional paper-and-pencil tests. The prescribed MOE English multiple assessment in primary schools is shown in Table 1 (Revised from Yeh, 2001; Chen, 2003).

Test-based assessment is generally a paper-and-pencil test or an oral test, often containing discrete-point and integrative tests (Yeh, 2001). This assessment focuses on students' ability to convey meanings for authentic purposes in an interactive context (Chen, 1999; Chen, 2001; Shih et al., 1999; Chen, 2003).

Task-based assessment is often teacher-prepared, involving teacher-student interaction, student-student interaction, body movement and teamwork. It includes role-plays, chanting, oral presentation, games, station running, problem-solving tasks, and group discussion (Mitchell & Parker, 2002).

Assessment by portfolio hinges on the evaluation of "a purposeful collection of students' work that demonstrates to students and others their efforts, progress, and achievements in given areas" (Genesee & Upshur, 1996, p. 99). Students'

Table 1

Types of multiple assessment

Types	Contents	Focus	Remarks
Summative assessment	Task-based assessment	English proficiency	1. Summative assessment accounts for about 30% of the final grades.
	Test-based assessment	English proficiency	2. Test-based tests refer to mid-term and final exams.
Formative assessment	Task-based assessment	English proficiency	1. Formative assessment accounts for about 70% of the final grades.
	Class observation	Learning attitudes and methods, cultural understanding	2. Test-based tests here refer to quizzes. Test-based tests should be held as less as possible.
	Assignment assessment	English proficiency, learning attitudes and methods, cultural understanding	
	Portfolio assessment		
	Test-based assessment	English proficiency	

portfolios provide a continuous record of language development, reflect the students' learning processes, demonstrate their real progress, give teachers and students opportunities to set goals and provide an opportunity for parents, students, and other teachers to view concrete results of learning.

Observation refers to the purposeful examination of teaching and/or learning events through systematic processes of data collection and analysis. Flippo (1997) specifies that the practice of observing students in class as they do their work, assignments, and other activities, in a natural, authentic manner, can provide the teacher valuable information regarding students' current and developing interests, motivations, strategies, and work habits.

Assessment and teacher beliefs

Generally beliefs are formed through personal experiences and interactions in daily and life and interpretations of events individuals have engaged in (Al-Sharafi, 1998; see Hsieh, 2002). These beliefs are transformed into attitudes, which in turn influence intentions, with intentions becoming the bases for decisions that lead to action (Bauch, 1984). In educational contexts, this concept of belief system governs teaching behaviors, with individual pedagogies reflecting a teacher's beliefs about language teaching (Bauch, 1984; Graves, 2000; Huang, 1997). Teachers' thought processes and instructional decisions are influenced by their beliefs (Borg, 1999). These beliefs largely direct teachers' choices and practice, such as addressing teaching objectives, designing lessons, selecting tasks and activities, and assessing student performance (Rios, 1996). Hence, in the classroom teachers not only impart knowledge to their students but

also consciously or unconsciously pass or impose their beliefs about learning on students (Horwitz, 1988). Cheng's (1997) findings revealed that teachers' beliefs about foreign language learning had crucial impacts on students' anxiety about foreign language learning, as Chinese teachers seemed to emphasize the importance of excellent pronunciation, immediate error correction, vocabulary memorization, and grammar rules. Because teachers' beliefs about language teaching are so influential and have such a great impact, studies on teachers' beliefs should be encouraged.

To date, the interconnection between teachers' beliefs and assessment within the context of recent developments in English language teaching in elementary schools in Taiwan has not been explored yet. It is this interconnection that motivated the present study aimed at exploring elementary school EFL teachers' beliefs and practices of multiple assessments. In particular, this study sought to answer the following questions:

1. What were elementary school EFL teachers' beliefs of multiple assessments?
2. What assessment types were frequently applied by elementary school EFL teachers?
3. Were there any discrepancies between EFL teachers' beliefs and their practices of multiple assessments? What factors contributed to these discrepancies?
4. Were EFL teachers' beliefs of multiple assessments influenced by their educational background, age, undergraduate degree major/concentration, training programs they received, and teaching experiences?
5. Were EFL teachers' practices of multiple assessments influenced by their educational background, age, undergraduate degree major/concentration, training programs they received, and teaching experiences?

Method

The study was conducted for one academic year, commencing in August 2006. There were 520 elementary school EFL teachers who participated in the study. Participants completed a questionnaire on their beliefs and practices of multiple assessment (The questionnaire consisted of three sections as shown in Table 2.

To establish validity of the questionnaire, selected professors from the field of multiple assessment and experienced EFL teachers in Taiwan were asked to review the questionnaire. A pilot study was conducted to test the reliability of the questionnaire. Data was analyzed using SPSS 13.0. The questions which were not statistically significant ($p < .05$) in the critical ratio test were deleted. Reliability analysis yielded a Cronbach's of .85, for the questionnaire items.

Data analyses used descriptive statistics, t-test, one-way ANOVA and bivariate correlation.

Demographic profile of EFL teachers

Of the total 520 EFL teachers, 503 (96.7%) were female and only 17 (3.3%) were male. Tables 3, 4, 5, 6, 7, and 8 present the demographic profile of the participants in this study.

Table 2

Questionnaire framework

Questions	Content		Question type
1–9	Personal information	Gender, age, educational background, undergraduate major, number of years of English teaching, graders EFL teachers are teaching, average class size, city/country of work, TESOL-related training program that teachers have taken	
10–19	EFL teachers' perceptions	1. Understanding of concepts of MA 2. Knowledge of key points of MA implementation	Fixed-alternative questions
35–37	EFL teachers' beliefs	Teachers' beliefs in assessment	Fixed-alternative questions
38–43		Teacher's beliefs in impact of assessment on teaching	
44–47		Teachers' beliefs in effects of assessment on students' learning	
48–49		Teachers' beliefs in parents' attitudes towards students' learning	
20–23	EFL teachers' practices, difficulties, and MA training needs	Assessment practices	Fixed-alternative questions
50–56		1. Purposes of assessment practices 2. Teachers' assessment practices 3. Difficulties of applying MA 4. Reasons for not using MA	Fixed-alternative questions and multi-multiple questions
33–34		Teachers' needs for MA workshops and training programs	Fixed-alternative questions
57	Teachers' opinions about MA		An open-ended question

Table 3

EFL teachers' age

Areas	Hsinchu County	Taoyuan County	Taipei County	Taipei City	Keelung City	Total
Age Range	Number					
21–30 years old	30	31	81	47	27	216
31–40 years old	35	32	55	56	27	205
41–50 years old	15	24	23	16	14	92
Above 50	1	1	2	1	2	7
Total	81	88	161	120	70	520

Table 4

EFL teachers' undergraduate major

Undergraduate majors	Number	Percent (%)	Cumulative Percent (%)
TESOL	80	15.4	15.4
Elementary education	211	40.6	56.0
Early childhood education	38	7.3	63.3
Chinese language and literature	67	12.9	76.2
Foreign language	103	19.8	96.0
Others	21	4.0	100.0
Total	520	100.0	

Table 5

Teaching experience

EFL teaching experience (in years)	Number	Percent (%)	Cumulative Percent (%)
Less than 1 year	16	3.1	3.1
1-2	86	16.5	19.6
3-5	201	38.7	58.3
6-10	186	35.8	94.0
More than 11 years	31	6.0	100.0
Total	520	100.0	

Table 6

Participants' professional TESOL-related training program

TESOL training program types	Number	Percent (%)	Cumulative Percent (%)
TESOL education	83	16.0	16.0
MOE's TESOL training program	176	33.8	49.8
20 credits TESOL training program	101	19.4	69.2
12 credits TESOL training program	59	11.3	80.6
8 credits TESOL training program	19	3.7	84.2
60 hours TESOL training program	48	9.2	93.5
30 hours TESOL training program	15	2.9	96.3
Others	19	3.7	100.0
Total	520	100.0	

Table 7

Average class size of participants

Class size	Number	Percent (%)	Cumulative Percent (%)
10 or below	7	1.3	1.3
11–15	15	2.9	4.2
16–20	22	4.2	8.5
21–25	71	13.7	22.1
More than 25	405	77.9	100.0
Total	520	100.0	

EFL teachers' perceptions, beliefs and practices of MA

The level of EFL teachers' perceptions of multiple assessments was based on respondents' scores on Factor Perceptions containing 10 question items (Q10–Q19). The scores were a sum of these 10 items, which ranged from a low of 10 to a high of 40. If the EFL teachers have a better understanding of the concepts of assessment and multiple assessments, and of the ways to assess students' English learning, the level of beliefs falls into the range of 32.5–40.

The level of EFL teachers' beliefs of multiple assessments was based on respondents' scores on factor beliefs containing 15 items (Q35–Q49). Scores ranging from a low of 15 to a high of 60 were a sum of these 15 items. If the EFL teachers have strong beliefs of multiple assessments, the level of beliefs falls into the range of 48.75–60.

In addition, the level of EFL teachers' practices of multiple assessments was based on respondents' scores on factor practices containing 15 items (Q20–Q34). Scores ranged from a low of 15 to a high of 60 were a sum of these 15 items. If the EFL teachers frequently use multiple assessments or alternative assessments in their teaching, the level of practices falls into the range of 48.75–60. Table 8 shows the descriptive statistics of (1) factor perceptions, (2) factor beliefs, and (3) factor practices of multiple assessments (MA).

The scores of EFL teachers' perceptions of multiple assessment were based on the 10 items on the four-point Likert scale of "strongly disagree", "disagree", "agree", and "strongly agree", ranging from 10 (with each item indicating "strongly disagree") to 40 (with each item indicating "strongly agree").

Table 8

EFL teachers' perceptions, beliefs, and practices of MA

Factor type	Mean	SD	Minimum	Maximum	Range	Skewness	Kurtosis
Perceptions	33.14	3.23	28	40	12	.466	-.651
Beliefs	49.32	4.14	41	59	18	.704	-.308
Practices	48.02	3.57	33	57	24	-.245	.325

N = 520

Table 9

Descriptive statistics of EFL teachers' perceptions of MA

Question items	SD (%)	D (%)	A (%)	SA (%)	Total
10. I understand the concept of portfolio assessment in instruction.	0 (0)	24 (4.6)	365 (70.2)	131 (25.2)	520
11. I understand how to guide students to collect materials for portfolio assessment.	8 (1.5)	55 (10.6)	323 (62.1)	134 (25.8)	520
12. Portfolio assessment improves students' self-assessment ability.	0 (0)	22 (4.2)	310 (59.6)	188 (36.2)	520
13. I understand how to use classroom observation (e.g., attitudes, behavior or performance) to improve instruction.	0 (0)	0 (0)	193 (37.1)	327 (62.9)	520
14. I understand how to implement a process of classroom observation.	0 (0)	0 (0)	209 (40.2)	311 (59.8)	520
15. I understand the concept of test-based assessment.	0 (0)	0 (0)	373 (71.7)	147 (28.3)	520
16. I understand how to design test-based assessments.	0 (0)	8 (1.5)	365 (70.2)	147 (28.3)	520
17. I understand the concept of task-based assessments (e.g., games, role play, and group discussion).	0 (0)	8 (1.5)	355 (68.3)	157 (30.2)	520
18. I understand how to design task-based assessments.	0 (0)	30 (5.8)	326 (62.7)	164 (31.5)	520
19. I am capable of setting the criteria to assess students' language performance.	0 (0)	38 (7.3)	356 (68.5)	126 (24.2)	520
Total	8 (0.15)	185 (3.56)	3175 (61.06)	1832 (35.23)	5200 (100.00)

EFL teachers' perceptions of MA

Table 9 shows that 96 % of the teachers reported that they have a clear understanding of the concept of MA. In response to Question 11, the majority knew how to guide students to collect material for portfolio assessment. Moreover, 95.4% of the EFL teachers reported that they knew the concept of portfolio assessment (Question 10). In addition, for Question 19, the majority agreed that they were capable of setting criteria to assess students' language performance.

EFL teachers' beliefs of MA

The majority of the respondents (98.5%) agreed that the ability to implement assessment is an important resource of a teacher. All the respondents reported that they disagree with the notion that traditional tests are more effective than alternative assessment of the four language skills. Concerning the tension between EFL teachers and students, 91.7% respondents disagreed with the idea of increased tension between students and teacher, resulting from multiple assessments.

EFL teachers also have a positive opinion and strong beliefs in the implementation of multiple assessments. They believe that use of multiple

assessment leads to a more focused instruction, increased reflection of teaching practices, diagnosis of students' learning difficulties, and planning for more engaging classroom activities. Data from answers to Questions 44–47 indicated that EFL teachers believed use of multiple assessments lead to the development of more practical scenarios for using English.

Table 10

Descriptive statistics of EFL teachers' beliefs of MA

Question items	SD (%)	D (%)	A (%)	SA (%)	Total
35. Assessment plays an important role in the teaching process.	0 (0)	0 (0)	132 (25.4)	388 (74.6)	520
36. The ability to implement assessment is an important capability for a teacher.	0 (0)	8 (1.5)	172 (33.1)	340 (65.4)	520
37. Traditional paper-and-pencil tests are more effective than alternative assessments in understanding students' listening, speaking, reading, and writing skills.	272 (52.3)	248 (47.7)	0 (0)	0 (0)	520
38. The tension between the teacher and his/her students will increase.	164 (31.5)	340 (65.4)	16 (3.1)	0 (0)	520
39. The teacher will focus more on students' performance capabilities.	0 (0)	0 (0)	412 (79.2)	108 (20.8)	520
40. The teacher will understand the effect of his/her teaching more easily.	0 (0)	0 (0)	341 (65.6)	179 (34.4)	520
41. It will be easier for the teacher to assess students' learning achievements in listening, speaking, reading and writing.	0 (0)	0 (0)	342 (65.8)	178 (34.2)	520
42. The teacher can easily figure out students' difficulties in learning.	0 (0)	0 (0)	349 (67.1)	171 (32.9)	520
43. The teacher will design more situations to make students active learners.	0 (0)	0 (0)	356 (68.5)	164 (31.5)	520
44. The assessment style can lower students' anxiety for assessments.	0 (0)	72 (13.8)	301 (57.9)	147 (28.3)	520
45. Students can understand more about their own learning problems.	0 (0)	8 (1.5)	411 (79.0)	101 (19.5)	520
46. Students can easily find out suitable approaches to their language learning.	0 (0)	71 (13.7)	379 (72.9)	70 (13.4)	520
47. Students will develop more practical skills to use English.	0 (0)	0 (0)	401 (77.1)	119 (22.9)	520
48. Parents will care more about students' performance than grades.	0 (0)	190 (36.5)	275 (52.9)	55 (10.6)	520
49. Parents will be better informed about their children's progress at schools.	0 (0)	8 (1.5)	347 (66.7)	165 (31.8)	520

Note: SD = Strongly Disagree; D = Disagree; A = Agree; SA = Strongly Agree; N = 520

EFL teachers' practices of MA

The majority of the respondents reported that they used task-based assessment more than traditional paper-and-pencil assessment (65.8%–agreed; 25.0%–strongly agreed). All the respondents reported that they integrated formative assessment into their teaching activities (Question 23). Moreover, 95.9% indicated that they used more alternative assessments than traditional tests (Question 24). Similarly, responses to Question 25 constituted 97.3% of EFL teachers who reported that they applied multiple assessments.

Lastly, concerning EFL teachers' need for workshops or programs related to multiple assessments, most of the respondents (87.7%) expressed the necessity to attend such events, supporting their strong perceptions of the salience of multiple assessments. Considering the recent inception of EFL in the elementary schools

Table 11
Descriptive statistics of EFL teachers' practices of MA

Question items	SD (%)	D (%)	A (%)	SA (%)	Total
20. I apply task-based tests more often than paper-and-pencil tests.	0 (0)	48 (9.2)	342 (65.8)	130 (25.0)	520
21. I emphasize students' language learning process more than their achievements.	0 (0)	80 (15.4)	355 (68.3)	85 (16.3)	520
22. In addition to grades, I give students written descriptions on their report cards.	0 (0)	150 (28.8)	262 (50.4)	108 (20.8)	520
23. I integrate formative assessment as a part of my teaching activities.	0 (0)	0 (0)	232 (44.6)	288 (55.4)	520
24. In my practice, I used more alternative assessments than traditional tests.	7 (1.3)	14 (2.7)	282 (54.2)	217 (41.7)	520
25. In my practice, I applied multiple assessments.	2 (0.4)	12 (2.3)	289 (55.6)	217 (41.7)	520
26. In my practice, I used more formative assessment than summative assessments.	2 (0.4)	12 (2.3)	289 (55.6)	217 (41.7)	520
27. In my practice, I used more task-based assessments than paper-and-pencil tests.	7 (1.3)	14 (2.7)	282 (54.2)	217 (41.7)	520
28. In my practice, I used a lot of classroom observations to examine students' learning.	2 (0.4)	12 (2.3)	291 (56.0)	215 (41.3)	520
29. In my practice, I spent a lot of time doing portfolio assessments.	31 (6.0)	114 (21.9)	294 (56.5)	81 (15.6)	520
30. When assessing students, I used a lot of teacher assessment.	0 (0)	0 (0)	150 (28.8)	370 (71.2)	520
31. When assessing students, I used a lot of peer assessment.	16 (3.1)	122 (23.5)	338 (65.0)	44 (8.5)	520
32. When assessing students, I used a lot of student self-assessment.	30 (5.8)	140 (26.9)	318 (61.2)	32 (6.2)	520
33. It is necessary to hold workshops or programs about the use of multiple assessments.	0 (0)	13 (2.5)	310 (59.6)	197 (37.9)	520
34. I need to attend workshops on the use of multiple assessments.	0 (0)	64 (12.3)	316 (60.8)	140 (26.9)	520

Note: SD = Strongly Disagree; D = Disagree; A = Agree; SA = Strongly Agree; N = 520

this intention to attend events promoting multiple assessment indicates a desire to develop the use of MA alongside classroom instruction.

The Pearson correlation was computed between EFL teachers' beliefs and practices, yielding a value of .186. The result showed that the relationship between beliefs and practices was positively significant ($p < .01$), which meant that the stronger beliefs on the multiple assessments the EFL teachers had, the more frequently they used multiple assessment in their teaching practices (see Table 12).

To test whether EFL teachers' perceptions of multiple assessments were related to their practices, Pearson correlation was computed. The analysis revealed that the relationship between perceptions and practices was not significant. This meant EFL teachers' understanding and notions of multiple assessment did not significantly determine whether they frequently used multiple assessment in their teaching practices (see Table 13).

Likewise, using the Pearson correlation between perceptions and beliefs was not significant (see Table 14). This indicated that EFL teachers' understanding and concepts of multiple assessment was not significantly correlated to their beliefs of multiple assessments.

Overall, only the relationship between beliefs and practices was positively significant.

Table 12

Correlation between EFL teachers' beliefs and practices of MA

	EFL teachers' beliefs	EFL teachers' practices
EFL teachers' beliefs	1	.186**
EFL teachers' practices	.186**	1

Note: ** $p < .01$

Table 13

Correlation between EFL teachers' perceptions and practices of MA

	EFL teachers' perceptions	EFL teachers' practices
EFL teachers' perceptions	1	.161
EFL teachers' practices	.161	1

Note: Two-tailed

Table 14

Correlation between EFL teachers' perceptions and beliefs of MA

	EFL teachers' perceptions	EFL teachers' beliefs
EFL teachers' perceptions	1	.363
EFL teachers' beliefs	.363	1

Note: Two-tailed

Relationship between beliefs/practices and age/experience of EFL teaching

EFL teachers in this study belonged to different age groups and possessed differing ESL teaching experience. One-way ANOVA tests reported the following relationships.

(a) The relationship between beliefs and age was not statistically significant ($F(3, 516) = 1.989, p = .115$). This meant that participants in different age groups did not differ significantly in their means of scores in beliefs of multiple assessments (see Table 15). The relationship between practices and age was not statistically significant ($F(3, 516) = 1.829, p = .141$), either. This revealed that participants in different age groups did not have significantly different mean scores in practices of multiple assessments.

(b) The relationship between beliefs and years of ESL teaching experience was statistically significant ($F(4, 515) = 12.614, p < .000$). This meant that participants in the five groups with different years of EFL teaching experience differed significantly in their means of scores in beliefs of multiple assessment (see Table 16). In addition, the results showed that relationship between EFL teaching experience and their practices was statistically significant ($F(4, 515) = 7.427, p < .000$). This revealed that participants in the five groups with different years of EFL teaching experience had significantly different mean scores in the practices of multiple assessments.

The significant omnibus F test was obtained in the one-way ANOVA test for both the relationship between EFL teachers' years of teaching and practices of

Table 15

ANOVAs of age and beliefs & age and practices

Types		Sum of squares	df	Mean square	F	Sig.
Beliefs	Between groups	101.737	3	33.912	1.989	.115
	Within groups	8795.986	516	17.046		
	Total	8897.723	519			
Practices	Between groups	69.437	3	23.146	1.829	.141
	Within groups	6529.407	516	12.654		
	Total	6598.844	519			

Table 16

ANOVAs of EFL teaching experience and beliefs & practices

Types		Sum of squares	df	Mean square	F	Sig.
Beliefs	Between groups	793.924	4	198.481	12.614	.000
	Within groups	8103.799	515	15.736		
	Total	8897.723	519			
Practices	Between groups	359.903	4	89.976	7.427	.000
	Within groups	6238.941	515	12.114		
	Total	6598.844	519			

Note: $p < .0125$

multiple assessments and the relationship between EFL teachers' years of English teaching and beliefs. It revealed that there was at least one pair of groups which differed significantly in means. To examine the location of group differences, the statistical procedures of post hoc multiple comparisons were applied as this study did not propose hypotheses about specific group differences. The Tukey test (aka Tukey HSD test) and the Scheffe test, were used¹.

(c) In terms of the relationship between EFL teachers' years of English teaching and beliefs, the results showed that in Tukey HSD post hoc multiple comparisons, there were 5 pairs in which the means were statistically different. They were Group 1 vs. Group 2, Group 1 vs. Group 3, Group 1 vs. Group 4, Group 3 vs. Group 4, and Group 4 vs. Group 5 (see Tables 17 and 18). In the conservative Scheffe test, there were 4 pairs whose means differed significantly. They were Group 1 (less than one year of EFL teaching experience) vs. Group 2 (1–2 years of EFL teaching experience), Group 1 (less than one year of EFL teaching experience) vs. Group

Table 17
Relationship between different years of ELT teaching and beliefs/practices

Variables	Years of EFL teaching M (SD)					F	Sig.	T**	S***
	Less than 1	1–2	3–5	6–10	More than 11				
	1	2	3	4	5				
Beliefs	45.00 (.00)	49.65 (3.89)	48.66 (3.85)	50.56 (4.52)	47.45 (1.12)	12.61	.000		
Practices	47.88 (2.60)	46.48 (3.82)	48.16 (3.63)	48.78 (3.33)	46.84 (2.63)	7.43*	.000*		

Note: * The mean difference was significant at the .0125 level.
 ** The locations of significant group differences in Tukey HSD post hoc multiple comparisons concerning beliefs: 2 > 1; 3 > 1; 4 > 1; 3 > 4; 5 > 4; concerning practices: 3 > 2; 4 > 2.
 *** The locations of significant group differences in Scheffe post hoc multiple comparisons concerning beliefs: 2 > 1; 4 > 1; 4 > 3; 5 > 4; concerning practices: 3 > 2; 4 > 2.

Table 18
Mean difference between years of teaching and beliefs in Tukey and Scheffe post hoc test

	Less than 1	1–2 years	3–5 years	6–10 years	More than 11
Less than 1		-4.65*	-3.66	-5.57*	
1–2 years	4.65*				
3–5 years	3.66			-1.91*	
6–10 years	5.57*		1.91*		3.12
More than 11				-3.12*	

Note: *The locations of significant group differences in both Tukey HSD and Scheffe post hoc multiple comparisons.

¹ In the post hoc multiple comparisons test of this study, since group sizes were unequal (16 for "less than 1 year", 86 for "1–2 years", 201 for "3–5 years", 186 for "6–10 years", and 31 for "more than 11 years"), harmonic mean sample size was used.

Table 19

Mean difference between years of teaching and practices in Tukey and Scheffe post hoc test

	Less than 1	1–2 years	3–5 years	6–10 years	More than 11
Less than 1					
1–2 years			-1.69*	-2.40*	
3–5 years		1.69*			
6–10 years		2.40*			
More than 11					

Note: *The locations of significant group differences in both Tukey HSD and Scheffe post hoc multiple comparisons.

4 (6–10 years of EFL teaching experience), Group 3 (3–5 years of EFL teaching experience) vs. Group 4 (6–10 years of EFL teaching experience), and Group 4 (6–10 years of EFL teaching experience) vs. Group 5 (more than 11 years of EFL teaching experience).

As to the relationship between EFL teachers' teaching experience and practices, the results revealed that in both Tukey HSD post hoc multiple comparisons and Scheffe test, there were two pairs in which the means were statistically different. They were Group 2 (EFL teachers who had 1–2 years of English teaching experience) vs. Group 3 (EFL teachers who had 3–5 years of English teaching experience) and Group 2 (EFL teachers who had 1–2 years of English teaching experience) vs. Group 4 (EFL teachers who had 6–10 years of English teaching experience) (Table 19).

Other findings on EFL teachers' practices of MA

Results showed that all the EFL teachers reported that the two most important purposes of assessment were to understand students' learning achievement and to understand students' progress. Nearly all the respondents thought that the purposes of assessment practices were to evaluate teaching effectiveness and to evaluate whether teaching activities achieve their objectives. These results revealed that for most of the respondents their assessment practices had multiple purposes.

Respondents reported that they preferred (a) alternative assessments (30%) and (b) both traditional and alternative (70%), while none chose traditional tests as the main assessment type. It was obvious that alternative assessments were the main choice, but traditional tests were not abandoned by some of the respondents.

Among the types of multiple assessments, results showed that paper-and-pencil tests were still used by the respondents as shown in Table 20.

Classroom observation was the second frequently-used assessment type. In Taiwan, English teaching at elementary schools provided students with two classes a week. Flippo (1997) specifies that the practice of observing students in class as they do their work, assignments, and other activities, in a natural, authentic

Table 20

Types of multiple assessments

Assessment type	Frequency	Percent (%)
1. Classroom observation	512	98.5
2. Paper-and-pencil tests	520	100.0
3. Portfolio assessment	219	42.1
4. Task-based assessment	474	91.2
5. Others	76	14.6

Table 21

Evaluators of multiple assessments

Evaluator	Frequency	Percent (%)
1. Teacher assessment	496	95.4
2. Peer assessment	274	52.7
3. Parent assessment	69	13.3
4. Student self-assessment	251	48.3
5. Others	0	0.0

manner, can provide the teacher valuable information regarding students' current and developing interests, motivations, strategies, and work habits.

The nature of assessment is presented in Table 21, with teachers as the main evaluator, followed by peers and self (Table 21).

In general, about fifty percent of EFL teachers used students as evaluators. Self-assessment provides students with opportunities to engage in checking their own assignment and assessment, while peer-assessment provides students with opportunities to assess other students. When employing these two methods, EFL teachers have to take into consideration the proficiency level of students and the contents of the assessment. Only when students reach the relevant ability, receive the necessary guidance and training, and have enough practice can they perform well in self-assessment. Using self-assessment is expected to be proposed in conjunction with other strategies and skills, and one of the recommended ways is to include students in the process through self and peer assessment. In terms of the time of assessment, 100% of the respondents reported that they used formative assessment, while 94.0% reported that they employed summative assessment. When EFL teachers apply alternatives to traditional assessment, they will choose different types of informal assessment tasks. Accordingly, formative assessment would be in use, too. As pointed out by Brown (2004, p. 6), "most of our classroom assessment is formative assessment: evaluating students in the process of 'forming' their competencies and skills with the goal of helping them to continue that growth process."

Of the 11 difficulty areas, time constraints with work overload contexts, large class size, and time-consuming activities mostly contributed to teachers' non-use of multiple assessment. It does seem that the "time" factor poses as a challenge

Table 22

Difficulties contributing to non-use of MA

Assessment time	Ranking	Frequency	Percent	Frequency	Percent
		Difficulties		Reasons	
1. Increased teachers' workload	3	369	71.0	369	71.0
2. Time-consuming	1	417	80.2	417	80.2
3. Subjective grading	4	339	65.2	339	65.2
4. Unfamiliar with MA techniques		86	16.5	86	16.5
5. Expensive		23	4.4	23	4.4
6. Difficult to grade		63	12.1	63	12.1
7. Too much to teach; without enough time	1	417	80.2	417	80.2
8. Too many students in my classes	1	417	80.2	417	80.2
9. Too many classes	2	409	78.7	409	78.7
10. Difficult to work with parents		15	2.9	15	2.9
11. Others		24	4.6	24	4.6

to implementing multiple assessment. As to the fourth difficulty, EFL teachers thought that multiple assessment was largely subjective.

Conclusions

Firstly, regarding EFL teachers' perceptions of multiple assessments, the study found that the EFL teachers had a better understanding of the concepts of assessment and multiple assessments. Most of them not only understood the concept of portfolio assessment but also knew how to use classroom observation. The study also found that EFL teachers had strong beliefs of multiple assessments. Most of the respondents believed that multiple assessments were more practical than the traditional paper-and-pencil tests. All of them believed that EFL teachers will find it easier to assess students' learning achievement in listening, speaking, reading, and writing. They believed that teachers can easily diagnose students' difficulties in learning. All of them believed that students will effectively develop strategies on using English if multiple assessments are used. In addition, results revealed that in general EFL teachers applied multiple assessments or used alternative assessment. The study found that many EFL teachers not only had strong beliefs of multiple assessments, but also applied most of the multiple assessments in their teaching. However, for some of the assessment types, especially those which time-related, EFL teachers preferred not to use them, such as portfolio assessment. For example, the percentage of participants who understood the concept of portfolio assessment was almost 20 times those who did not understand it. Besides, most of the respondents agreed or strongly agreed that portfolio assessment could improve students' self-assessment ability. However, when it came to the practice of portfolio, the

percentage of participants who applied portfolio assessment (72.1%) was just about 2.58 times those who did not applied portfolio assessment (27.9%). This means there was a discrepancy between perception/beliefs and practices.

Secondly, the correlation coefficient ($r = .186$) between EFL teachers' beliefs and practices showed that their relationship was positively significant at the level of .01, which meant the stronger beliefs on the multiple assessments the EFL teachers had, the more frequently they used multiple assessments in their teaching practices.

Thirdly, the one-way ANOVA tests showed that the relationship between beliefs and years of ESL teaching experience was statistically significant ($F(4, 515) = 12.614, p < .000$). This meant that participants in the 5 groups with differing EFL teaching experiences varied significantly in their beliefs of multiple assessment. In addition, the results showed that the relationship between EFL teaching experience and their practices was statistically significant ($F(4, 515) = 12.114, p < .000$). This revealed that participants in the 5 groups with differing years of EFL teaching experience had significantly different mean scores in the practices of multiple assessments.

Fourthly, no respondent choose traditional tests as the primary assessment. Multiple assessments were the main choice, but traditional tests were not abandoned by quite a few respondents. As a result, a combination of traditional tests and alternative assessments received the highest percentage. Concerning the types of multiple assessments, the study found that the paper-and-pencil test type was the most frequently-used assessment type. That means EFL teachers in Taiwan are confident with this assessment type. The second frequently-used assessment was classroom observation.

Finally, in terms of the difficulties of employing multiple assessments and the reasons why EFL teachers would not like to use multiple assessments, of the 11 difficulty options, the highly challenging ones were time constraints with work overload contexts, large class size, time-consuming activities, and notions of subjectivity surrounding multiple assessment.

Overall, the findings reveal the evolving teaching and assessment scenarios in some schools in Northern Taiwan, with teachers of differing experiences and backgrounds using multiple assessment alongside traditional assessment. The results suggest some positive implications for reconsidering class size, teacher workload, and training on multiple assessment to promote the use of multiple assessment in the elementary curriculum.

CORRESPONDENCE

Any correspondence should be directed to Yu-Ching Chan, National Taipei University of Education (yuching@tea.ntue.edu.tw).

References

- Alderson, J., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Al-sharafi, A. (1998). An investigation of the beliefs and practice of foreign language teachers: A case study of five American high school foreign language teachers in Leon County. Unpublished doctoral dissertation. College of Education of Florida State University.
- Bailey, K. (1998). *Learning about language assessment: Dilemmas, decisions and directions*. Boston: Heinle & Heinle.
- Bailey, K.M. (2001). Observation. In R. Carter & D. Nunan (eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp.114-119). Cambridge: Cambridge University Press.
- Bauch, P. (1984). The impact of teachers' instructional beliefs on their teaching: Implications for research and practice. *ERIC Digest*. ED252954
- Borg, S. (1999). Teachers' theories in grammar teaching. *ELT Journal*, 53(3), 157-167.
- Brady, P. (1997). Assessment and the curriculum. In C. Cullingford (ed.), *Assessment versus evaluation* (pp. 8-23). London: Cassell.
- Brindley, G. (2001). Assessment. In R. Carter & D. Nunan (eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp.137-143). Cambridge: Cambridge University Press.
- Brown, J. (1989). Language program evaluation: A synthesis of existing possibilities. In R. Johnson (ed.), *The second language curriculum* (pp. 222-241). Cambridge: Cambridge University Press.
- Brown, H. (2001). *Teaching by principles*. New York: Longman.
- Brown, H. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Brown, J., & Hudson, T. (1998). *New ways of classroom assessment*. Alexandria, VA: Teachers of English to Speakers of Other languages.
- Bunn, S. (2001). Portfolio assessment in Texas. *Caves English Teaching*, 30, 28-31.
- Burnett, G. (1993). The assessment and placement of language minority students. *ERIC Digest*. ED 357131.
- Burns, P., Roe, B., & Smith, S. (2002). *Teaching reading in today's elementary schools*. Boston, MA: Houghton Mufflin.
- Cai, B. (1998). On teaching assessments in open education. In W. Deng (ed.), *Multiple assessments in open education* (pp. 20-34). Kaohsiung: Fu-Wen.
- Chan, Y., & Wu, G. (2004). A study of foreign language anxiety of EFL elementary school students in Taipei County. *Journal of National Taipei Teachers College*, 17(2), 287-320.
- Chan, Y., Chen, L., Chen, K., & Yu, V. (2003). The study of the practice of English carnival in elementary schools in Taipei City and Taipei County. *The proceedings of the 2003 conference of education and academic research on Taiwan's children* (pp.1-28). Taipei: National Taipei Teachers College.
- Chang, C.B. (1996). Multiple teaching assessment. *Workshop communication*, 13(3), 1-9.
- Chang, D.T. (1984). The study of the implementation of teaching assessments in junior high school. Unpublished master thesis. Taipei: National Taiwan Normal University.
- Chen, C., & Ruan, S. (2001). New outlook of elementary school English assessment—Assessments through running stops. *Proceedings of the Eighteenth International Symposium on English Teaching* (pp. 497-510). Taipei: Crane.
- Chen, C. (1999). Vivid and active learning assessments. In W. Dai (ed.), *Teaching methods and materials for elementary school students* (pp. 222-255). Taipei: Crane.
- Chen, H. (2003). A study of primary school English teachers' beliefs and practices in multiple assessments: A case study in Taipei City. Unpublished master theses. Taipei: National Taipei Teachers College.
- Chen, L. (2000). Portfolios in second/foreign language classroom assessment. *English teaching and learning*, 25(2), 28-47.
- Cheng, M. (1997). The impacts of teachers' beliefs on students' anxiety about foreign language learning. *The proceedings of the Sixth International Symposium on English Teaching* (pp. 113-129). Taipei: Crane.
- Chern, C., Ruan, S., & Yeh, S. (2001). *Activity-based assessment for children and elementary school students*. Taipei: Ur Best.

- Cohen, A. (2001). Second language assessment. In M. Celce-Murcia (ed.), *Teaching English as a second or foreign language* (pp. 515-534). Boston: Heinle & Heinle.
- Cole, D. (2000). *Portfolios across the curriculum and beyond*. Thousand Oaks, Calif.: Crown Press.
- Fischer, C., & King, R. (1995). *Authentic assessment: A guide to implementation*. Thousand Oaks, Calif.: Crown Press.
- Flippo, R. (1997). *Reading assessment and instruction: A qualitative approach to diagnosis*. Florida: Harcourt Brace.
- Gattulo, F. (2000). Formative assessment in ELT primary (elementary) classrooms: An Italian case study. *Language Testing*, 17(2), 278-288.
- Genesee, F. (2001). Evaluation. In R. Carter & D. Nunan (eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp.144-150). Cambridge: Cambridge University Press.
- Genesee, F., & Upshur, J. (1996). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.
- Glaser, R. (1990). *Testing and assessment: O tempora! O mores!* Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.
- Graves, K. (2000). *Designing language courses: A guide for teachers*. Boston: Heinle & Heinle.
- Gronlund, N. (1998). *Assessment of student achievement* (6th ed). Boston: Allyn and Bacon.
- Guba, E., & Lincoln, Y. (1981). *Effective evaluation*. San Francisco: Jossey-Bass Publishers.
- Halliwell, S. (1992). *Teaching English in the primary classroom*. New York: Longman.
- Hansen, J. (1992). Literacy portfolios: Helping students know themselves. *Educational Leadership*, 49, 66-68.
- Hasselgren, A. (2000). The assessment of the English ability of young learners in Norwegian schools: An innovative approach. *Language Testing*, 17(2), 261-277.
- Heaton, J. (1990). *Writing English language tests*. New York: Longman Inc.
- Hewitt, G. (1995). *A portfolio primer: Teaching, collecting, and assessing students writing*. Portsmouth, NH: Heinemann.
- Horwitz, E. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal*, 72, 283-294.
- Hsieh, H. (2002). Teachers' beliefs about English learning: A case study of elementary school English teachers in Taipei County. Unpublished master thesis. Taipei: National Taipei Teachers' College.
- Hsieh, Y.F., Lu, C.Y., & Yeh, H.N. (2000). Implementation of portfolio assessment in a sixth grade EFL classroom. *The proceedings of 2000 national workshop and conference on the application of English teaching* (pp. 112-139). Taipei: Crane.
- Hsu, Y. (2003). Researching multiple assessment: Evidence from EFL elementary school classroom. *English Teaching and Learning*, 28(1), 29-60.
- Huang, C. (2001). *Authentic assessment*. Taipei: Experimental Elementary School of National Taipei Teachers' College.
- Huang, S.L. (1997). The domestical situations and prospection of the study of teachers' beliefs. *Journal of Humanity and Society of National Chung-Hsing University*, 6, 135-152.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Jiang, W. (2000). The crucial abilities for teachers' instruction assessment. *A Tribune for Elementary Education*, 137, 3-10.
- Johnstone, R. (2000). Context-sensitive assessment of modern language in primary (elementary) and early secondary education: Scotland and the European experience. *Language Testing*, 17(2), 123-143.
- Lynch, B. (1996). *Language program evaluation: Theory and practice*. Cambridge: Cambridge University Press.
- Madsen, H. (1983). *Techniques in testing*. Oxford: Oxford University Press.
- Meiselles, E. (1995). Teachers as participants in the National Reading Examinations. In P. Owen & P. Pumfrey (eds.), *Curriculum and assessment issues: Messages for teachers* (pp. 155-161). London: The Farmer Press.
- Meisels, S. (2000). On the side of the child. *Young Children*, 55(6), 16-19.
- Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass Publishers.

- Mitchell, H., & Parker, S. (2002). *Pop up 1: Special edition*. London: MM Publications.
- Moya, S., & O'Malley, J. (1994). A portfolio assessment model for ESL. *The Journal of Educational Issues of Language Minority Students*, 13, 13-36.
- Neuman, S., Copple, C., & Bredekamp, S. (2000). *Learning to read and write: Developmentally appropriate practices for young children*. Washington, DC: National Association for the Education of Young Children.
- Newman, C., & Smolen, L. (1993). Portfolio assessment in our schools: Implementation, advantage, and concerns. *Mid-Western Education Researcher*, 6, 28-32.
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.
- O'Malley, J., & Pierce, L. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. White Plains, NY: Addison-Wesley.
- Paris, S. (1994). *Becoming reflective students and teachers*. Washington, DC: American Psychological Association.
- Pierce, B. (1992). Demystifying the TOEFL reading test. *TESOL Quarterly*, 26(4), 665-689.
- Pierce, L., & O'Malley, J. (1992). *Performance and portfolio assessment for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Popham, W.J. (1995). *Classroom assessment*. Needham Heights, MA: Allyn & Bacon.
- Ren, C. (2001). Design of multiple assessments. *Guidance of Elementary Education*, 40(5), 49-56.
- Richards, J. (2002). 30 years of TEFL/TESL: A personal reflection. *Proceedings of the Eleventh International Symposium on English Teaching/Fourth Pan Asian Conference* (pp. 141-161). Taipei: Crane.
- Richards, J., & Lockhart, C. (1994). *Reflective teaching in second language classroom*. New York: Cambridge University Press.
- Rios, F. (1996). *Teacher thinking in cultural contexts*. New York: State University of New York Press.
- Scott, W., & Ytrebert, L. (1990). *Teaching English to children* (2nd ed.). New York: Longman.
- Shih, Y. (2001). Evaluation of the MOE primary school English teacher training program. *English Teaching and Learning*, 26(1), 86-107.
- Shih, Y.H., Chou, C.T., Chen, S.C., Chu, H.M., & Chen, C.Y. (1999). *The model of English teaching and assessment in the Nine-Year Integrated Curriculum*. Ministry of Education, R.O.C.
- Shohamy, E. (1994). The role of language tests in the construction and validation of second-language acquisition theories. In E. Tarone, S. Gass, & A. Cohen (eds.), *Research methodology in second-language acquisition* (pp. 133-142). Hillsdale, NJ: Lawrence Erlbaum.
- Shohamy, E. (1998). Applying a multiplism approach. In E. Li, & G. James (eds.), *Testing and evaluation in second language education* (pp. 96-105). Hong Kong: Language Centre, Hong Kong University of Science and Technology.
- Shohamy, E. (2000). The relationship between language testing and second language acquisition revisited. *System*, 28(4), 541-553.
- Skehan, P. (1998a). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (1998b). Task-based instruction. In W. Grabe (ed.), *Annual Review of Applied Linguistics*, 1998 (pp. 134-135). New York: Cambridge University Press. .
- Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1, 185-211.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49, 93-120.
- Smith, P. (1995). A framework for literacy assessment. In P. Owen. & P. Pumfrey (eds.), *Curriculum and assessment issues: Messages for teachers* (pp. 121-134). London: The Falmer Press.
- Sutherland, J. (1999). Types of evaluation for primary school English. Paper presented at the *Conference on the Training of Elementary School EFL Teachers* held at the National Taiwan Normal University, Ping Tung, June 11-13.
- Tannenbaum, J. (1996). Practical ideas on alternative assessment for ESL students. *ERIC Digest*. ED395500.
- Weir, C. J. (1990). *Communicative language testing*. Prentice Hall.
- Wishon, P., Crabtree, K., & Jones, M. (1998). *Curriculum for the primary years: An integrative approach*. Upper Saddle River, NJ: Merrill, An Imprint of Prentice Hall.

- Wolf, D., Bixby, J. III, & Gardener, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of Research in Education, 17*, 31-74.
- Worthen, B.R. (1993). Critical issues that will determine the future of alternative assessment. *Phi Delta Kappa, 74*(6), 444-54.
- Yeh, H. (2001). The implementation of portfolio assessment in elementary school. *Caves English Language Teaching, 30*, 11-14.
- Yin, R. (1994). *Case study research: Design and methods* (2nd ed). London: Sage.
- You, C. (2000). Instruction in small size class and multiple assessments. *Humanities and Social Sciences Communication, 11*(4), 53-63.
- Zahorik, J. (1995). Evaluating. In L. Anderson (ed.). *International encyclopedia of teaching and teacher education* (pp. 283-286). Cambridge: Cambridge University Press.

Appendix 1

A Questionnaire on "A Study of Elementary School English Teachers' Beliefs in and Practice of Multiple Assessments in Northern Taiwan" (adapted from Yang (2000), Chen (2003), and Chan (2006).

Dear Teacher,

This is a questionnaire for academic study about elementary school English teachers' implementation of multiple assessments in the Northern Taiwan. The purpose of this study is to understand your beliefs and practices of multiple assessments. Your answers to the question items in this questionnaire are very valuable and important. They will be used exclusively for the academic study. In this questionnaire, except for some marked questions which may be given more than one answer, most of the questions require only one appropriate answer. Please answer each question based on your own personal situations. Thank you for your cooperation and assistance.

Daniel Yu-ching Chan
Ph.D./Associate Prof.
Graduate School of Children English Education
National Taipei University of Education
August 20, 2006

Part I. Background Information (Please mark your answer in the appropriate box.)

1. Gender
 Male Female
2. Age
 21-30 31-40 41-50 above 51
3. Educational background (Please indicate the highest education level completed)
 Junior college Undergraduate Master's degree Doctorate
4. Undergraduate major
 TESOL Elementary education Early childhood education
 Chinese language and literature Foreign language Others
5. English teaching experience
 less than 1 year 1-2 years 3-5 years 6-10 years
 more than 11 years
6. The grade(s) that you are teaching (check all that apply)
 Grade 1 Grade 2 Grade 3 Grade 4
 Grade 5 Grade 6
7. The average number of students in each class which you are teaching this year:
 10 or below 11~15 16~20 21~25
 more than 25
8. Which city or county are you teaching now?
 Hsinchu County Taoyuan County Taipei County Taipei City
 Keelung City
9. What kind of training program related to TESOL have you taken?
 TESOL education MOE's ESL training program
 20 credits TESOL training program 12 credits TESOL training program
 8 credits TESOL training program 60 hours TESOL training program
 30 hours TESOL training program Others

Part II. Perception, Beliefs and Practices.

(For questions 10–49 please mark the appropriate box based on the four options: 4 = Strongly agree; 3 = Agree; 2 = Disagree; 1 = Strongly disagree)

	1	2	3	4
10. I understand the concept of portfolio assessment in instruction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. I understand how to guide students to collect materials for portfolio assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Portfolio assessment improves students' self-assessment ability.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. I understand how to use classroom observation (e.g. attitudes, behaviors or performances) to improve instruction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. I understand how to implement a process of classroom observations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. I understand the concept of test-based assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. I understand how to design test-based assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. I understand the concept of task-based assessments (e.g. games, role play, and group discussion).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. I understand how to design task-based assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. I am capable of setting the criteria to assess students' language performance.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. I apply task-based tests more often than paper-and-pencil tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. I emphasize students' language learning process more than their achievements.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. In addition to grades, I give students written descriptions on their report cards.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. I integrate formative assessment as a part of my teaching activities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. In my practice, I used more alternative assessments than traditional tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. In my practice, I applied multiple assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. In my practice, I used more formative assessment than summative assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. In my practice, I used more task-based assessments than paper-and-pencil tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. In my practice, I used a lot of classroom observations to examine students' learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. In my practice, I spent a lot of time doing portfolio assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. When assessing students, I used a lot of teacher assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. When assessing students, I used a lot of peer assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. When assessing students, I used a lot of student self-assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33. It is necessary hold workshops or programs about the use of multiple assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34. I need to attend workshops on the use of multiple assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35. Assessment plays an important role in the teaching process.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36. The ability to implement assessments is an important capability for a teacher.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
37. Traditional paper-and-pencil tests are more effective than alternative assessments in understanding students' listening, speaking, reading, and writing skills.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- | | 1 | 2 | 3 | 4 |
|--|--------------------------|--------------------------|--------------------------|--------------------------|
| In my opinion, when multiple assessments are used by a teacher, | | | | |
| 38. the tension between the teacher and his/her students will increase | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 39. the teacher will focus more on students' performance capabilities. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 40. the teacher will understand the effect of his/her teaching more easily. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 41. it will be easier for the teacher to assess students' learning achievements in listening, speaking, reading and writing. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 42. the teacher can easily figure out students' difficulties in learning. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 43. the teacher will design more situations to make students active learners. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 44. the assessment style can lower students' anxiety for assessments. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 45. students can understand more about their own learning problems. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 46. students can easily find out suitable approaches to their language learning. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 47. students will develop more practical skills to use English. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 48. parents will care more about students' performance than grades. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 49. parents will be better informed about their children's progress at schools. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
50. The purposes of my assessment practices were to: *(check all that apply)*
- Understand students' learning achievement.
 - Find out more about students' learning difficulties.
 - Make use of results from assessment as a basis for remedial teaching.
 - Evaluate whether teaching activities achieve their objectives.
 - Understand students' progress.
 - Evaluate teaching effectiveness.
 - Other
51. When assessing students, I used: *(check one box)*
- Traditional tests
 - Alternative assessments
 - Traditional tests and alternative assessments
52. When implementing multiple assessments, I used: *(check all that apply)*
- Classroom observation
 - Paper-and-pencil tests
 - Portfolio assessment
 - Task-based assessment
 - Other
53. In terms of the evaluator, when implementing multiple assessments, I used:
(check all that apply)
- Teacher assessment
 - Peer assessment
 - Parental assessment
 - Student self-assessment
 - Other
54. In terms of the time of assessment, when implementing assessments, I used:
(check all that apply)
- Formative assessment (i.e., assess students during class)
 - Summative assessment (i.e., assess students in a period of time, e.g., monthly exam, midterm exam, final exam)
 - Quizzes
 - Other

55. What were the difficulties I encountered when I used multiple assessments?

(check all that apply)

- Multiple assessments increased teacher's workload.
- Multiple assessments were time-consuming.
- The grading of multiple assessments was subjective.
- I was unfamiliar with how to use the implementing techniques of multiple assessments.
- Multiple assessments were expensive.
- It was difficult to grade with multiple assessments.
- I had too much to teach and didn't have the time to use multiple assessments.
- There were too many students in my classes.
- I had too many classes of students to teach.
- It was difficult to work with parents in multiple assessments.
- Other

56. The reasons I don't like to use multiple assessment: *(check all that apply)*

- Multiple assessments increase teacher's workload.
- Multiple assessments are time-consuming.
- The grading of multiple assessments is subjective.
- I am unfamiliar with how to use the implementing techniques of multiple assessments.
- Multiple assessments are expensive.
- It is difficult to grade with multiple assessments.
- I have too much to teach and do not have the time to use multiple assessments.
- There are too many students in my classes.
- I have too many classes of students to teach.
- It is difficult to work with parents in multiple assessments.
- Other

57. What is your opinion about multiple assessments?

Thank you very much for your assistance!