English Language Teachers' View on Assessment and their Reflection on Teaching Practices: A Mongolian Case Study

Jargaltuya Ragchaa, PhD Candidate
University of Szeged, Hungary
Dornod University, Mongolia

Doi: 10.19044/ejes.v6no2a1 URL:http://dx.doi.org/10.19044/ejes.v6no2a1

Abstract

In Mongolia, teachers' attitudes towards large-scale assessments have been a largely uncovered area. Therefore, this paper focuses on examining the general overview of English language teachers' belief about state-level assessments, and their reflection on teaching practices. Participants of this study were 307 teachers of primary and secondary school teachers, and 36 of them were English language teachers of Dornod province in Mongolia. Independent sample t-test was used to explore how English teachers change their instructions in teaching English compared to other subjects teachers. Result showed that they usually search for more effective teaching methods, take less liberty on how they design their lessons, reduce instructional content, and focus more on Educational standards. As a result of a correlation analysis, English language teachers' assessment view is significantly related to the content of the assessment that is designed by teachers in a class. Based on the results, it can be concluded that teachers focus more on the assessment content that they design for progress and final exams. They prefer to prepare students for this assessment by making them practice the test items that are similar to the school achievement test items during classes. Understanding the reasons for ineffective instructions can help policy makers and teachers to change the assessment content and its accountability, and would also help to improve their classroom instructions to have better learning outcome.

Keywords: Large-scale assessments, test-based accountability, instructional change, English language teaching.

Introduction

A lot of children are learning English language in different schools around the world. English is increasingly perceived as a basic competence to succeed in life. Mongolia has adopted English language as a second language,

and schools are offering English as a main mandatory subject. English is included as one of the main subjects in school achievement tests. Majority of parents often search for schools that offers English Language and that have better quality programs for their children.

Given the recognized importance of English, language education and its assessment are changing from developing students' academic skills to the use of English in real life. Nikolov (2016) noted that one of the best programs of English, documented by recent interest, is content and language integrated learning. Johnstone (2009) and Rixon (2013, 2016) remark that this new development poses new opportunities and challenges for assessment. Nikolov (2016) added that this shift towards assessment and accountability are not limited to foreign language programs. However, there is an international trend in educational assessment for accountability in public education policies in all subjects and competencies. Assessment and its accountability have become inseparable parts of education and, based on the assessment, program accountability calls for the quality of education to be continually improved. However, recent studies indicate that in most cases, assessment is administrated to see that the implementation of standards and curriculum are being met. Based on the results of the study, assessment can be used for ranking the schools, teachers, and students in a bid to improve the teaching and learning process (Nikolov, 2016). The aim of this paper is to specify and understand what English language teachers think of state-level assessments and their usefulness, and how their instructions and test preparation strategies are changed due to their perceptions of state-level assessments.

Literature Review

Early research indicates different directions of the impacts of high stakes tests. Thus, they have both negative (anxiety and fear) and positive (changes of teaching instruction and test taking strategy) effects on learning and teaching practices. External pressure can lead teachers to critically revise their practices and adapt effective teaching strategies (Terhart, 2013). In contrast, Hamilton et al. (2002) argued that test-based accountability can also lead to negative reallocation of instructional time to focus on tested aspects of the standards to the exclusion of untested aspects of the standards. English language instructors are encouraged (Baker & Westrup, 2000) to use many methods to teach receptive skills in pre-stages and post-stages. On the other hand, Alkaff (2013) noted that students concentrate more on terminology and that they are usually tested with multiple choice questions because of limited practice on everyday interactions in the classroom.

Tran (2012) highlighted the importance of validity, reliability, practicality, equivalency, authenticity, and wash back of second language assessment. He explained that test validity needs to measure the test takers'

real ability based on empirical and theoretical research. Bachman and Palmer (1996 cited in Tran, 2012) say that reliability refers to similar results when the test is administered on different occasions. Practicality refers to the relationship between the resources (human and materials, time, and location) and the use of the test. Equivalency and authenticity indicate whether or not the test is directly based on curriculum standards or instructional activities. Brown and Hudson (1998 cited in Tran, 2012) pointed out that a wash back is the reflection of testing and assessment on the language teaching curriculum and instruction. These studies show that including all of these criteria for writing tests is really important to assess students' actual skills and their learning outcome. Second language testing assesses learners' progress and their specific skills. Therefore, language instructors need to design tests to measure the learners' functional use of language, not a specific linguistic point.

Consequently, the most important thing test makers need to consider in language assessment is to understand the roles of abilities and contexts, the interactions between them, and the influence of ability and context on the performance of language assessment tasks (Fox et al., 2007). Powers (2010) observed that language receptive (reading and listening) and productive skills (speaking and writing) are assessed differently. Receptive skills are usually assessed through computer-based and paper-pencil with multiple choice items, while productive skills are assessed with performance-based tests. Language testing experts and language researchers such as Hakuta and Beatty (2000), Bailey and Butler (2003), and Garcia, McKoon and August (2006) have criticized previous English language assessments used for ESL students. This is because those assessments do not measure up with the development of the academic English language skills that students need to become successful in a school settings. Language educators noted that an interactional approach was becoming more important in language teaching and assessment. For example, Bachman (2007) and Chapelle (1998) noted that the English language program includes skill-based, trait-based, task-based, and interactional approaches in a given context. Chapelle (1998) remarks that an interactional approach to language learning improves communicative language abilities. Chalhoub-Deville (2003) noted that language competence is a process involving improvement over time in combining knowledge and context with language performance.

Across the world, English teachers have different assessment views. Language assessments can be different or similar in different countries. Rixon (2013) found that, at the end of primary school years, English language assessments were different in some countries. For instance, in France, at the end of primary school years, teachers complete an evaluation which covers five skills areas : (1) listening comprehension, (2) oral interaction, (3)

individual speaking with no interaction (e.g. reproducing a model, a song, a rhyme, a phrase, reading aloud, giving a short presentation), (4) reading comprehension, and (5) writing. In Taiwan, instructors are now developing their own English proficiency tests (Rixon, 2013) at the primary school level. The purpose of their proficiency test is to assess the effectiveness of English instruction and to identify those in need of remedial teaching. In Finland, many primary schools use a voluntary "national" test of English designed by the English teachers' association of Finland to guide their final grading of students and to get some information for them about how they are doing against the average of other schools (Rixon, 2013). A New National Curriculum in Cyprus was implemented in September 2011. It introduced English at the primary level, emphasized the role of portfolio assessment, and introduced content and language integrated learning (Rixon, 2013).

Teachers' Instructional Change based on Assessment and Accountability

Researchers differently indicate that high stakes tests contribute to negative (anxiety and fear) and positive (changes of teaching instruction and test taking strategy) effects in teaching practices. A good dose of pressure can force teachers to adapt effective teaching strategies (Terhart, 2013). Tóth and Csapó (2011) explored how Hungarian teachers in elementary schools felt pressured by different stakeholders than their counterparts in upper secondary schools. However, they claimed that greater incentives and heightened external pressure were needed to induce school agents to raise educational quality. Hamilton et al. (2005) noted that the integration of mechanisms of educational accountability system can positively affect the quality of education. As they reported, the mechanisms—incentives, information, and assistance—are likely to affect student achievement primarily by altering what occurs in the classroom: Incentives are intended to motivate teachers to focus on the goals embodied in the standards, information from tests should provide data to guide instructional decision making, and assistance should help them improve their practice" (Hamilton et al., 2005, p.3).

According to Hamilton et al. (2002), test-based accountability can lead educators to work harder and to adopt better curricula or more-effective teaching methods. It can lead to coaching students to perform better by focusing on aspects of the test that are relevant to the domain the test is intended to represent. Due to a test-based accountability system, teachers may pay more attention to test-taking strategies. Often, multiple-choice state school achievement tests differ widely from the format used in classroom tests. Pederson and Yager (2014 in Ngang, Hong & Chanya, 2014, p.536) remarked that becoming a highly qualified teacher in today's educational system is dependent on how well teachers work together with their principals and colleagues. Through collective work, teachers explore the potential to practice

more effective decision making as a skill for supporting acquisition of additional professional knowledge and skills.

A number of other studies have shown that test-based accountability programs have had a positive impact on students' test scores (e.g. Carnoy & Loeb, 2002; Jacob, 2005; Linn & Dunbar, 1990; Nichols, Glass, & Berliner, 2012).

In contrast, Boyd et al. (2008 in Fuller & Ladd, 2012, p.13) noted that teachers avoid high stakes tests that may induce teachers' anxiety of unwanted inquisition, loss of flexibility in classroom practices, a feeling of coercion to teach based on the test, and fear for their jobs. Tóth and Csapó (2011) found that in Hungary, teacher beliefs' about changes in their teaching are rather similar in elementary and lower secondary schools. However, the level of agreement in the case of many of the statements differs between elementary and upper secondary school teachers. In Hungary, teachers typically refuse to narrow down the curriculum due to the national assessment system; they focus their efforts more on students with poor results in the state tests by giving extra-curricular tutoring.

Koretz et al. (2001) found that test-based accountability has no effect, or even a negative effect on students' knowledge and skills. Hamilton et al. (2002) points out that test-based accountability can also lead to negative reallocation of instructional time to focus on tested aspects of the standards to the exclusion of untested aspects of the standards. In addition, high-stakes testing may become a barrier to the development of intrinsic motivation as its implementation is generally accompanied by a high amount of pressure on students and teachers (Moore & Waltman, 2007). Thus, the various studies reviewed above show the usefulness of test-based accountability systems. Herman and Golan (1991) noted that high-stakes testing leads to a narrowing of curricula and instruction, and such testing appears to influence teaching and learning within schools. Teachers spend most of their time and attention to increase students' test scores rather than focus on student learning. Thus, state test results, under conditions of accountability pressure, remain a critical issue to understand when designing and implementing accountability measures. Meaningful learning requires a critical approach based on the productive use of assessment in stimulating educational reform.

The NBETPP (National Board on Educational Testing and Public Policy) (2003) reports that teachers often spend more time on subjects that are tested with high stakes, and less time on non-tested subjects. Therefore, students have limited time to practice with fine arts, physical education, foreign language, and other extra-curricular activities. Similarly, Abrams, Pedulla and Madaus (2003) and Abrams (2004) conducted a survey among Florida teachers and the result showed that teachers had reallocated instructional schedules, allowing for more time to be spent on tested content

while reducing the time for the material that would not appear on the large scale assessment. Hence, they reduce the time spent on fostering activities in order to prepare students for the state test.

Hadley (2010) carried out a survey on 12 school principals from eight different district schools in the state of Utah to explore their opinions about how high-stakes testing impacts teaching and learning. The findings showed that principals were concerned that teachers should teach a curriculum that would result in improved test scores. Additionally, the principals encouraged teachers to use the results of large scale assessments to guide their instruction to produce high test scores. Eslami-Rasekh and Valizadeh (2008) conducted a survey on Iranian young EFL teachers. They responded that they felt more successful in applying instructional strategies than in managing an EFL class. They also reported that their ability to motivate and engage students to learn English was not as high as their ability to use instructional strategies.

Teachers' Test Preparation Strategies based on Assessment and Accountability

Clearly, educational researchers should pay attention to teachers' test preparation strategies caused by large-scale assessment and accountability systems. In the NBETPP report (2003), teachers responded to some questions related to preparing their students for the state-level test such as test preparation methods and amount of time spent on test preparation. They stated that more time is spent due to high-stakes tests with intense preparation using materials that closely resemble the test. Also, they try to motivate their students to do well in the state test.

In addition, majority of teachers changed their assessment practices by modeling their own classroom tests following the format of the state test. Abrams et al. (2003) report that teachers from high-stakes states spend more time than do their counterparts in low-stakes states preparing students for the state test. Abrams (2004) also found that in Florida, many teachers and schools are highly stressed by the pressure to improve student test performance. Sixty-three percent of teachers indicated that the pressure was so much that they had little time to teach anything that would not appear on the test. Furthermore, majority of them reported that they found ways to raise test scores without improving learning. Hadley (2010) remarks that test subjects and test preparation activities restrict the amount of time spent on a particular subject, and the tests dictated the kind of teaching strategies used, resulting in fewer activities, less creativity, and less benefit to the students.

Methodology

Participants

The participants were 262 different subject teachers and 36 English language teachers from 19 schools in Dornod province. Dornod province lies at the eastern part of Mongolia and includes a major city, Choibalsan. Those 19 schools were in Choibalsan and in nearby villages (soums) in the surrounding metropolitan area. The subjects were 100% female with a mean age of 33.8 and a mean teaching experience of 9.4 years. 75% of them were Bachelor of Arts holders and 25% of them were MA degree holders.

Instruments

The teachers' view on educational assessment and accountability questionnaire was created based on numerous international questionnaires (Hamilton, Berend, & Stecher, 2005; Moore & Waltman, 2007). This is the questionnaire of the IPEA (International Project for the Study of Educational Accountability Systems) project. Tóth and Csapó (2011) adapted this questionnaire to fit into the Hungarian context. In adapting this version into Mongolian context, some questions related to International assessment were discarded because Mongolia is not included in some International studies such as PISA and TIMMS. The questionnaire consisted of seven blocks of questions (61 items). Each block represented a particular assessment or accountability procedure: (1) Teachers' background questions consisting of five items, (2) views on large-scale assessments consisting of 10 items, (3) test preparation strategies consisting of nine items, (4) perceived pressure for different types of assessment consisting of seven items, (5) amount of test practice consisting of three categorical items, (6) changes in instructional practice including 20 statements, and (7) perceived pressure from different stakeholders consisting of seven items. Teachers' opinions were assessed on a four point Likert scale (1=disagree; 4=agree).

Results

An independent sample t-test was used based on the acceptance of the large scale assessments to compare the means of selected 36 English teachers' opinions with other teachers' opinions. Table 1 shows that English language teachers in Dornod province have similar ideas compared to other teachers. Specifically, they think that school achievement tests should be conducted on a regular basis (M=3.6, SD=.60), that tests contribute to an increased effort in schools (M=3.4, SD=.93), that tests provide an objective basis to evaluate schools (M=3.3, SD=.79), and that these tests are important for work in schools (M=3.2, SD=.98). They also, like teachers of other subjects, somewhat disagree with the view that school achievement tests support the debate about the concept of competence (M=2.2, SD=.96), and they provide a basis for

discussion among colleagues (M=2.3, SD=.96). Other teachers, however, had different opinions about the statements "School achievement tests are important for work in schools" (t=-1.16, p<.05) and "School achievement tests are not useful for my job as a teacher" (t=-1.46, p<.05).

Table 1. English language teachers' view on school achievement tests

	School achievement tests	Groups	N	M	SD	t	P
1	Should be conducted on a regular basis	ENG	36	3.6	.60	-.60	n.s
		Other	262	3.6	.65		
2	Contribute to an increased effort in schools	ENG	34	3.4	.93	.24	n.s
		Other	253	3.3	.85		
3	Provide an objective basis to evaluate schools	ENG	35	3.3	.79	1.70	n.s
		Other	257	3.0	1.0		
4	Are important for work in schools	ENG	36	3.2	.98	-1.16	p<.05
		Other	255	3.3	.79		
5	Create more problems than solutions	ENG	35	2.7	.93	-.82	n.s
		Other	257	2.8	.92		
6	Provide a basis for discussion among colleagues	ENG	34	2.3	1.0	-1.79	n.s
		Other	252	2.6	1.1		
7	Support the debate about the concept of competence	ENG	36	2.2	.96	-.69	n.s
		Other	253	2.4	1.0		
8	Are barely applicable for individual student evaluations	ENG	34	2.1	1.1	-.29	n.s
		Other	257	2.6	1.0		
9	Only cause trouble in schools	ENG	36	2.0	.96	5.50	n.s
		Other	253	2.0	1.0		
10	Are not useful for my job as a teacher	ENG	36	1.5	.80	-1.46	p<.05
		Other	257	1.8	1.0		

Note: N= number of participants, M= mean value of participants, SD= standard deviation, t= t-value (the size of the difference between means), p= p-value (significance level), n.s=not significant.

An independent sample t-test was also used to explore which instructional changes were mostly made by English teachers in teaching English in comparison to other teachers. The results in Table 2 below show that English language teachers and other subjects teachers usually search for more effective teaching methods (M=3.8, SD=.35), take less liberty on how they design their lessons (M=3.7, SD=.62), reduce instructional content (M=3.6, SD=.60), and focus more on Educational standards (M=3.5, SD=.56). English teachers have different opinions on the statement "I search for more effective teaching methods" (t=1.32, p<.05) and "I have narrowed down the curricular content of my instruction" (t=-2.02, p<.05) compared to teachers of other subjects in general. The results suggest that English language teachers are less willing to narrow down their curricular content than teachers in other

fields, and that they spend more effort searching for effective teaching methods.

Table 2. English language teachers' instructional changes

	School achievement tests	Groups	N	M	SD	T	P
1	I search for more effective teaching methods	ENG	35	3.8	.35	1.32	p<.05
		Other	246	3.7	.47		
2	I take less liberties on how I design my lessons	ENG	36	3.6	.62	.33	n.s
		Other	244	3.6	.57		
3	I reduce instructional content	ENG	36	3.5	.60	1.34	n.s
		Other	244	3.3	.74		
4	I focus more on Educational standards	ENG	36	3.5	.56	.11	n.s
		Other	247	3.5	.67		
5	My instruction focuses more strongly on competences rather than content	ENG	35	3.5	.78	1.61	n.s
		Other	247	3.5	.62		
6	I focus more strongly on multiple choice tests	ENG	36	3.4	.69	1.69	n.s
		Other	246	3.1	.83		
7	I focus more strongly on overarching competences (writing and reading in mathematics instruction)	ENG	36	3.3	.75	-1.22	n.s
		Other	239	3.4	.70		
8	I have narrowed down the curricular content of my instruction	ENG	36	3.0	.58	-2.02	p<.05
		Other	245	3.2	.72		

Note: N= number of participants, M= mean value of participants, SD= standard deviation, t= t-value (the size of the difference between means), p= p-value (significance level), n.s=not significant.

Based on confirmatory factor analysis, the following factors of changes in teachers' instructional practice were identified: (1) giving homework, (2) teaching methods, (3) content of the instruction, (4) testing strategy, and (5) teachers' attention to special students. A correlation analysis was done to identify how English teachers' view on school achievement tests (SATs) are related to their teaching practices. As a result, English language teachers' assessment view was found to be significantly related to the content of the instruction (r=.357, p<.05). The analysis also shows that teaching methods and testing strategy are correlated (r=.356, p<.05), and the content of the instruction is correlated with testing strategy (r=.334, p<.05) and attention to special students (r=.478, p<.01). These results are presented in Table 3 which follows.

Table 3. Relationship between assessment attitude and teachers' instructional changes

	Giving homework	Teaching methods	Content of the instruction	Testing strategy	Attention to special students
View on assessments					
Giving homework	.027				
Teaching methods	.137	-.133			
Content of the instruction	.357*	.058	.254		
Testing strategy	.031	-.074	.356*	.334*	
Attention to special students	.189	.037	.175	.478**	-.008

Note. * p<.05, ** p<.01.

English language teachers often use tasks in regular instruction that are similar to those in school achievement tests (M=3.6, SD=.47), discuss general task-taking strategies with students (M=3.6, SD=.58), have students practice test formats that are used in school achievement tests (M=3.5, SD=.66), and seek to improve students' motivation to do well on SATs (M=3.4, SD=.73). One significantly different statement compared to other teachers' answers was "I discuss general task-taking strategies with students" (t=1.49, p<.05). Other teachers, however, see more coherence between instructional content and tasks in SATs (t=-1.40, p<.05) and try to improve students' test taking skills (practice on public release tasks that are used in SATs) (t=-.90, p<.05) as summarised in Table 4 below.

Table 4. English language teachers' test preparation strategies

	Test preparation strategies	Groups	N	M	SD	T	P
1	I more often use tasks in regular instruction that are similar to those in school achievement test	ENG	34	3.6	.47	.21	n.s
		Other	253	3.6	.64		
2	I discuss general task-taking strategies with students	ENG	34	3.6	.58	1.49	p<.05
		Other	256	3.5	.81		
3	I practice test formats that are used in school achievement test	ENG	35	3.5	.66	-.55	n.s
		Other	255	3.5	.64		
4	I seek to improve students' motivation to do well on SATs	ENG	36	3.3	.72	-.52	n.s
		Other	258	3.4	.73		
5	I see to it that coherence between instructional content and the tasks of the SAT is increased	ENG	35	3.2	1.0	-1.40	p<.05
		Other	252	3.5	.74		
6	I try to improve students' test taking skills (practice on public release tasks that are used in SATs)	ENG	35	3.2	1.1	-.90	p<.05
		Other	260	3.4	.79		
7		ENG	35	2.1	1.0	-1.15	n.s

I set aside or put less emphasis, in regular instruction, on content that will not be tested	Other	259	2.3	1.0

Note: N= number of participants, M= mean value of participants, SD= standard deviation, t= t-value (the size of the difference between means), p= p-value (significance level), n.s=not significant.

Conclusion

Since English language is an important subject included in school achievement tests, it is important that English language teachers should believe that state-level assessments are important for their work and that the results of the assessment are linked with the school and teachers' efforts. The state-level assessments cause English teachers to focus more on the assessment content and influence their design of progress and final exams. English teachers also prefer to prepare students for the assessments by practicing test items that are similar to the school achievement test items during class. The results of this study will help give insights into the issues behind the teaching and learning process of English language education in Mongolia. It is important to explore the reasons behind ineffective teaching and learning strategies and their effect on learning achievement, and how English language instruction has been changing due to the educational assessment and accountability system in Mongolia. An independent sample t-test was used for exploring the frequencies and differences between the perceptions of assessment and accountability, and their instructional changes. The main results indicated that English language teachers think state-level assessments are important for improving the quality of language education since English language is included in school achievement tests. They also think it is better to conduct these assessments regularly.

However, they believe that these assessments are aimed only at evaluating schools, not for developing individuals' learning outcomes. English language teachers try to use more effective teaching methodologies even though they already do not have enough time to prepare their lessons due to their work load and the different types of assessments. Therefore, they reduce their instructional content and focus more on preparing students for exams. In addition, their view on the importance of large scale assessments influences the content of the assessment that they design for progress and final tests in their classes. Thus, this may be the reason why English teachers prefer to ask students to practice on the test formats that are used in the school achievement tests during class.

References:

1. Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, 42*(1), 8–29.
2. Abrams, L. M. (2004). Teachers' Views on High-Stakes Testing: Implications for the Classroom, Arizona State University Education Policy Studies Laboratory, Working Paper EPSL0401-104-EPRU.
3. Alkaff, A.A. (2013). Students' Attitudes and Perceptions towards Learning English, *AWEJ*, Vol.4. pp.106-121.
4. Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice: Designing and developing useful language tests, Oxford England: Oxford University Press.
5. Bachman, L. F. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessments. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E.Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 41–71). Ottawa, Canada: Ottawa University Press.
6. Bailey, A. L., & Butler, F. A. (2003). *An evidentiary framework for operationalizing academic language for broad application to K–12 education*: A design document (CSE Technical Report No. 611). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). *classes with few resources*. London: Continuum.
7. Baker, J. & Westrup, H. (2000). *The English language teacher's handbook: How to teach large classes with few resources*. London: Continuum.
8. Boyd, D., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., McDonald, M., Reininger, M., Ronfeldt, M., & Wyckoff, J. (2008). Surveying the landscape of teacher education in New York city: Constrained variation and the challenge of innovation. *Education Evaluation and Policy Analysis*, 30(4), 319-343.
9. Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. TESOL Quarterly, 32 (4), 653-675.
10. Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis.*Educational Evaluation and Policy Analysis*, *24*(4), 305–331.
11. Chalhoub-Deville, M. (2003). Second language interaction: Current perspectives and future trends, *Language Testing, 20*, 369–383.
12. Chapelle, C. A. (1998). Construct definition and validity inquiry in SLA research. In L. F. Bachman & A. D. Cohen (Eds.), *Second language acquisition and language testing interfaces* (pp. 32–70). Cambridge, England: Cambridge University Press.

13. Eslami-Rasekh, Z. & Valizadeh, K. (2008). Teachers' sense of self-efficacy, English proficiency, and instructional strategies: A study of nonnative EFL teachers in Iran. TESL EJ, 11(4), 1-19. Retrieved October 9, 2008 from http://tesl-ej.org/ej44/a1.pdf.
14. Fox, J., Wesche, M., Bayliss, D., Cheng, L., Turner, C.E., & Doe, C. (2007). *Language testing reconsidered*, University of Ottawa Press, Ottawa.
15. Fuller, S. C. & Ladd, H. F. (2012).*School Based Accountability and the Distribution of Teacher Quality Among Grades in Elementary Schools.* Panel Session for the Association for Education Finance and Policy, Boston, MA, March.
16. Garcia, G. E., McKoon, G., & August, D. (2006). Synthesis: Language and literacy assessment. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language minority children and youth* (pp. 269–318). Mahwah, NJ: Lawrence Erlbaum.
17. Hamilton, L. S., Berends, M., & Stecher, B. (2005). Teachers' Responses to Standards-Based Accountability. Working Paper. RAND Corporation, Santa Monica.
18. Hamilton, L, S., Stecher, B. M., & Klein, S. P. (2002). *Making Sense of Test-Based Accountability in Education.* RAND Corporation, Santa Monica.
19. Hakuta, K., & Beatty, A. (2000).Testing English-language learners in U.S. schools. Washington, DC: National Academies Press.
20. Herman, J., & Golan, S. (1991). Effects of Standardized Testing on Teachers and Learning—Another Look. Los Angeles, CA: National Center for Research on Evaluation, Standards and Student Testing, UCLA.
21. Jacob, B. A. (2005). Accountability incentives and behavior: The impact of high stakes testing in the Chicago public schools. *Journal of Public Economics*, *89*(5–6), 297-327.
22. Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In Koda, K. (1996). *L2 word recognition research: A critical review. Modern Language Journal*, 80, 14, 450–460.
23. Koretz, D., Mccaffrey, D., & Hamilton, L. (2001). *Toward a Framework for Validating GainsUnder High-Stakes Conditions*. CSE Technical Report.No. 551. Los Angeles, CA: Center for the Study of Evaluation, Universityof California.
24. Linn, R. L., & Dunbar, S. B. (1990). The Nation's report card goes home: Good news and bad about trends in achievement. *Phi Delta Kappan, 72* (2), October, 127–133.

25. Moore, J. L. & Waltman, K. (2007).*Pressure to Increase Test Scores in Reaction to NCLB: An Investigation of Related Factors*, Annual Meeting for the American Educational Research and Evaluation Association.
26. NBETPP Report (2003). Perceived effects of state-mandated testing programs on teaching and learning, Joseph J. Pedulla, Lisa M. Abrams, George F. Madaus, Michael K. Russell, Miguel A. Ramos, and Jing Miao Lynch, School of Education, Boston College.
27. Ngang, T. K., Hong, C. S., & Chanya, A. (2014). Collective work of novice teachers in changing teaching practices, *Procedia-Social and Behavioral Sciences*, 116 (21), pp. 536-540.
28. Nichols, S., Glass, G., & Berliner, D. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Education Policy Analysis Archives, 20* (20), 1–35.
29. Nikolov, M. (2016). A framework for young EFL learners' diagnostic assessment: Can dostatements and task types. In M. Nikolov (Ed.), *Assessing young learners' of English: Global and local perspectives* (pp. 65-92). New York: Springer. 15
30. Powers, D. E. (2010). The case for a comprehensive, four skills assessment of English language proficiency.*TOEIC Compendium study,*12.pp. *2-11*.
31. Rixon, S. (2013).*British Council Survey of policy and practice in primary English languag teaching worldwide*. London: British Council.
32. Rixon, S. (2016). Do developments in assessment represent the 'coming of age' of young learners English language teaching initiatives? The international picture. In M. Nikolov (Ed.), *Assessing young learners' of English: Global and local perspectives* (pp. 19-41). New York: Springer.
33. Terhart, E. (2013). Teacher resistance against school reform: reflecting an inconvenient truth, *School leadership management*, 33(5), 486-500.
34. Tóth, E. & Csapó, B. (2011). Comparing primary and high-school teachers' attitudes towards testing and accountability. Paper presented at the 14th European Conference for the Research on Learning and Instruction. Exeter, United Kingdom, August 30- September 3, 2011. 57.
35. Tran, T. H. (2012). Second Language Assessment for Classroom Teachers , Paper presented at MIDTESOL-2012, Ames, Iowa Rolla, Missouri, USA