

The Journal of Teaching Language Skills (JTLS)
3(2), Summer 2011, Ser. 63/4
(Previously *Journal of Social Sciences & Humanities*)

Iranian EFL Teachers' Views on Measuring Reading Comprehension

K. Sadeghi

Assistant Professor
Urmia University, Urmia
email: k.sadeghi@urmia.ac.ir

Abstract

Understanding the opinions of major role players in education (i.e. teachers, students, and policy makers) on all aspects of learning and teaching is influential to the success of the educational process. The purpose of this study was to investigate EFL (English as a Foreign Language) teachers' views on measuring reading comprehension. A cohort of 50 Iranian EFL teachers teaching at university, school and institute levels were invited to fill out a tailor-made questionnaire. Based on the analysis of 23 returned questionnaires, it was revealed that in most cases teachers did not have the freedom to make their own tests, or had to follow out-dated guidelines by relevant authorities.

Keywords: teachers' views, EFL, measurement, reading comprehension, Iranian, reading tests

1. Introduction

In an educational system, testing is an integral part of teaching. While in summative assessment, testing usually follows teaching and is the final loop of the chain, in formative assessment, testing comes in the middle of teaching and it is quite legitimate to think of testing which precedes teaching as in placement examinations. Whatever the order, it can be strongly argued that as far as educational process is concerned, teaching and testing are inseparable. While the major stakeholders in a testing task are the candidates who take the test, their ideas on what a test should be like usually counts the least primarily because they are non-professionals. (See Sadeghi 2008, as an example of the judgemental validity of cloze procedure from the standpoint of test-takers themselves).

Test construction process, particularly in the case of standardized and high-stakes tests, has mainly been researched from the perspective of test-makers who are in most cases experts in measurement. In educational settings, the burden of making a test is usually placed on teachers themselves who, despite having passed a course on language testing, are not usually conversant enough with the principles of language testing. In a country like Iran, English teachers are usually responsible for testing what they teach and for making their own tailor-made tests, particularly at school and university settings. Even when state-wide final-term achievement tests at school level or for University Entrance Exam are constructed, they are mainly constructed by teachers rather than professional test-making bodies. While making such tests, teachers usually use the guidelines provided by relevant educational authorities, which are sometimes sound, sometimes faulty, and sometimes outdated.

The fact that teachers are the very testers in educational systems warrants an investigation of their beliefs, priorities, biases and concerns about the measurement instruments they use or are instructed to construct. It was accordingly the aim of this study to look at teacher conceptions of tests of reading comprehension at school, university and institute levels. The nature of this study is largely qualitative so no

specific research questions or hypotheses were posed to be answered or tested. The major focus of this piece of research was an attempt to perceive the perceptions of Iranian EFL teachers about testing reading comprehension with a particular emphasis on the tests they have been using themselves at a stage of their teaching. Finding this out may not be such an easy task as it may seem since, as Buck, Ritter, Jensen and Rose (2010) rightly emphasize, "it is hard to know what teachers *really* think about testing" and "the attitudes of teachers toward testing is somewhat puzzling" (p. 50). Although there is much relevant previous research in the area of teachers' views on different aspects of education, not much has been said or written on their beliefs on the assessment side of the coin. Accordingly, a brief review of the relevant literature is presented next before attending to methodological issues.

2. Literature of Review

Investigating teachers' views has always been an important consideration in learning-teaching settings. Teachers' views have been looked at from different angles and on different aspects of education. A very small sample of such research during recent years is presented here. Gialamas and Nikolopoulou (2010), for example, compared pre-service and in-service Greek early childhood teachers' ideas on integrating ICT in their classes. Ranta (2010) looked at the ideas of 34 non-native teachers of English in Finnish secondary schools on teaching targets and found that Finnish teachers of English were aware of the lingua franca role of English and welcomed diversity. Chikasanda, Ortel-Cass and Jones (in press) investigated Malawian teachers' conceptions towards technical education and found that the teachers' views were shaped not only by their expectations and beliefs about the nature of technical education but by their perceptions about what students could gain from such education. Eret and Ok (in press) studied the views of 278 prospective English teachers studying at different state universities in Ankara and found that these would-be teachers had positive views on the quality of their instructors, the

curriculum implemented as well as the administrative support they received from their departments. However, they were generally unhappy about the physical environment of the departments where they studied.

Stevenson (2009) did a similar study on the relationship between teachers' beliefs and educational goals and subscribing to or adapting the mandated curriculum. The results indicated that Adaptor participants tended to have broader scores in terms of educational purposes than Subscribers. In a longitudinal case study, Levin and Nevo (2009) studied the changing beliefs of 10 elementary school teachers and found that their views had changed from pure status to multiple complementary visions which usually overlap.

In an attempt to find teachers' attitudes towards the value of educational research, Beycioglu, Ozer and Ugurlu (2010) found that 68% of the participating teachers valued educational research by seriously considering it since they were first qualified as teachers. In a similar study, Vanderline and van Braak (2010) investigated teachers', school leaders' and researchers' views on the gap between educational research and practice and found that the gap may be closed by establishing "professional learning communities" (p. 299). Elementary school teachers' views on critical pedagogy were studied by Yilmaz (2009). The researcher found that there was agreement between teachers on the principles of critical pedagogy, and while no significant difference was found between their ideas as far as gender was concerned, their views differed considerably from one another on variables such as educational background, professional seniority and the environment of the school where they worked.

As far as assessment is concerned, however, there is scant literature available when it comes to teachers' conceptions of testing in general and of measuring EFL reading comprehension in particular. Wang, Kao and Lin (2010), for instance, studied pre-service elementary teachers' views on the assessment of science learning and found that their assessment beliefs were coherent with a traditional view of

learning and that their performance mode of assessment was not properly developed. Troudi, Coombe and Al-Hamliy (2009) compared EFL teachers' views from Kuwait and the United Arab Emirates on assessment procedures in higher education and found that teachers' beliefs on assessment reflected the socio-political conditions governing their employment and that the managerial approaches to education heavily influenced the flexibility of teachers in exercising different assessment modes.

Studying kindergarten teachers' views on curriculum, instruction and assessment in the United Arab Emirates, Al-Momani, Ihmeideh and Momani (2008) found that teachers' assessment and instruction practices mainly focused on teaching and testing academic skills using direct instructional approaches. Comparing students' and teachers' views on using portfolios for assessment and learning among midwifery students, Mitchell (1994) observed that while students had negative feelings towards portfolios as assessment and teaching tools, their tutors showed awareness of the merits of these instructional and assessment tools.

In an attempt to understand teachers' views on testing programmes, Abrams, Pedulla and Madaus (2003) conducted a nation-wide survey and found that teachers reported that the pressures on them to raise test scores motivated them to attend more to learning and testing strategies that were dependent on the content and format of the final test with a negative washback on teaching.

Buck et al.'s (2010) interview results with 42 American teachers in five Arkansas schools showed some very interesting ideas on exams and testing. While their own literature review of the opinions of teachers on tests and testing published over the past five years in 3 education journals had revealed that "articles critical to testing outnumbered the favourable articles by an overwhelming 9- to- 1 ratio" (p. 50) and that "testing, in and of itself, is a negative process, as argued in some 90% of the literature" (p. 53), their own data showed the following five useful qualities for tests: Tests provide useful

information; Testing and standards help create a road map for the year's instruction; Test-preparation does not necessarily sap creativity for teachers or students; Testing can lead to collaboration; and Accountability is useful (pp. 51-53). Having briefly reviewed the current literature on teacher's views on education and assessment, we now turn to our own study.

3. Method

3.1 Participants

In order to understand the viewpoints of Iranian EFL teachers on the validity of the reading tests they constructed or used, three classes of teachers were identified out of the total population of EFL teachers in West Azarbaijan province of Iran. A convenience sample of 50 EFL teachers participating in this study came from practicing teachers at Urmia University, the Iran Language Institute, the Jahade Daneshgahi Language Centre and the Ministry of Education. The fact that teachers serving at the Ministry of Education, the Iran Language Institute and the Jahade Danesghgahi across Iran use the same teaching materials and that the teaching/learning activities and the syllabi followed are more or less the same imply that apart from the only teacher from the university sector (whose data was excluded from the final analysis), other teachers can be regarded as partially representative of their pertinent populations. The findings, however, may not be readily generalizable as discussed below since an opportunity or convenience sampling (Dörnyei, 2007) was employed for the selection of the participants. The following table shows the characteristics of the teachers whose data were included in the final analysis as unfortunately not all 50 teachers returned the questionnaires handed to them. Out of all questionnaires only 23 were returned.

Table 1: Characteristics of the participants

Characteristics <i>Questionnaire groups</i>	n	sex				Years of teaching		University degree
		M	F	mean	SD	mean	SD	
		University Teachers	1	0	1	34	---	
Secondary School Teachers	15	12	3	33	5.71	9.57	6.24	BA: 13 MA: 2
Private Language Centre Teachers	7	5	2	38.71	8.04	10.86	7.4	BA: 5 BSc: 1 MA: 1

Several reasons may have influenced less than expected degree of co-operation on the part of teachers, such as the coincidence of this research with final-term exams (which requires teachers to prepare and mark exam papers), not having enough time, being unwilling to participate, not valuing research, and the like.

3.2 Instruments and procedure

A questionnaire was constructed and reviewed several times before being used to elicit data on how English language teachers tested their students' reading comprehension, what they thought cloze tests were appropriate for, how they thought reading comprehension was tested or could be tested, and their attitudes towards tests of reading comprehension. The questionnaire was piloted on a few English teachers. Necessary revisions were made to make problematic items easier to understand, and it was also decided to allow teachers to give their answers to the questions either in English or Persian as piloting indicated that some teachers had difficulty in expressing themselves in English. No reliability index was estimated as the questionnaire was intended to provide qualitative rather than quantitative information.

Information was also asked about their name (optional), sex, age, the number of years they taught EFL, the institution/school/university in which they taught, the level at which they taught, and the university degree they held. (The questionnaire used in the study appears in the Appendix). 50 such questionnaires along with cover letters were distributed among English teachers teaching EFL at different levels (secondary school, pre-university, university, and private language centres) in 3 cities of West Azarbaijan province in Iran.

Questions 1-5 dealt with teachers' conception of reading comprehension, factors affecting reading comprehension, and similarities and differences between academic and real-life reading. These questions were intended as introductory questions to introduce the issue of how reading comprehension was tested. As only the answers to questions 6-17 are needed to know EFL teachers' viewpoints on tests of reading comprehension (with a particular focus on cloze testing), only the answers given to these latter questions are analyzed and discussed next. Teachers' ideas when quoted are italicised with long quotations indented.

It is also worth noting that initially all the questionnaires were expected to be answered in English as EFL teachers were thought to be proficient enough to provide their answers in English. When the researcher talked with a few respondents (especially at school level) about the fewer than expected number of returned questionnaires, they felt that some teachers might have had problems with the language needed to answer the questions and that had they been asked to answer either in English or in Persian, the proportion of returned questionnaires could have been greater. Based on such a suggestion, the researcher administered the same questionnaire to a few more EFL teachers, asking them to provide their answers either in English or in Persian. As such, two questionnaires were answered in Persian, one half in English, half in Persian, and all the rest in English.

4. Results

To give a better coherence and organisation to the content of the analysis, the relevant answers to each question in the questionnaire are brought separately and pertinent issues are accordingly discussed next. It is worth mentioning that the data produced by the only university teacher is excluded from the analysis (or only sporadically referred to), as the relevant ideas may in no way be representative of the general attitude of university teachers. Since not much variation was observed between two other cohorts of teachers' (i.e. school versus institute teachers) ideas, and since they shared more or less the same viewpoints and concerns, their answers to relevant questions are dealt with in the same place, rather than being categorized into different sections.

How do you think one's real-life comprehension of a passage is tested?
Most of the teachers seemed to have somehow misinterpreted the question. While the question asked how real-life reading *is tested*, most of the teachers talked about the way it *can be tested*. One teacher only seems to have correctly grasped the question: "*if he/she reads with ease or pleasure, this can show his/her degree of understanding*," meaning that real-life reading is not tested formally in a way academic reading is. Another teacher gave a similar answer: "*it is difficult to test because the readers' enjoyment is interrupted*". The rest mentioned how real-life reading *can be tested* using the following techniques: asking questions and interviewing; measuring performance based on reading; asking general questions after reading an article from mass media; asking to read a text from a newspaper and then "*analysing it correctly*"; using oral and written questions; giving a summary of the text; using (True/False) T/F, (multiple choice) m/c, yes/no (y/n), and wh-questions; and finally through cloze.

*How do you think one's academic reading comprehension is tested?
Which one do you consider a proper measure of comprehension of a given text? Why?*

Three teachers believed that cloze tests are good for such a purpose, but at advanced levels. Most of the teachers thought that some kinds of questions should be given after a text. Here are the techniques that teachers thought could be used/are used for testing academic reading comprehension: asking about text purpose; reading comprehension questions and performance based on reading; answering general questions after a carefully selected article; m/c, T/F questions, synonyms/antonyms, inferential and objective questions; and giving a summary after reading the text (oral or written). Some of the answers were unrelated, however, like "*by taking exams about it and pre-reading activities, then a summary*".

The second part of the question was understood differently by different teachers. Some had compared testing in academic reading with real-life reading. Some others compared different techniques they mentioned for testing academic reading. In the former group, while some believed that testing in real-life reading was a proper measure of reading comprehension, others believed that academic testing was better at least for students. The latter group felt that wh-questions, m/c questions, and cloze were better than other techniques in this regard. The reasons they brought were: they measured learners' full knowledge, they were easy for most students, and they measured language skills and reading comprehension in a real situation. One teacher noted that he did not know which technique properly measured reading comprehension, and therefore, suggested the use of more than one technique.

What methods do you use for testing reading comprehension of your students?

The following techniques were mentioned by respondents for measuring their students' reading comprehension: m/c questions, T/F questions, completion questions, and cloze; short-answer questions; general questions; summarising; written or spoken questions; explaining through one's own simple words; acceptable-scoring cloze;

wh-questions; y/n questions; and long-answer questions. One teacher stated that "*cloze is popular*" for such a purpose, and another teacher reported to have used cloze tests for 3rd and 4th year secondary school students only (relatively advanced school students). One teacher clearly indicated that he had to adhere to the rules and regulations set out by the Ministry of Education as how to test reading comprehension: "*As a rule ... I use sentence-comprehension (mini-comprehension) method and also reading 'unseen' texts*" [in which both comprehension and mechanics are tested].

Based on the researcher's own experience as a school teacher some 10 years ago, such 'mini-comprehension' sentences are taken from the comprehension passages students have already studied during English courses, and their true validity for measuring comprehension is well under question, especially because sometimes the choices given are even more difficult to understand than the stem of the item itself.

What do you try to test when you are testing reading comprehension? (for example, literal meaning, general meaning, details, inferential meaning, etc.)

The majority of the teachers regarded the main purpose of reading as gaining a general meaning, which was what they focused on during testing. Some teachers believed that depending on the purpose of the text and the exam both general meaning and details could be tested, while a few of them mentioned that all meaning types mentioned in the body of the question were important in constructing reading comprehension tests. One teacher looked at reading comprehension as reading 'skill' and stated that he would focus on both micro- and macro-skills in testing. Another teacher mentioned that he would also test "*memory, experience, intuition, mediation, and imagination*" in his reading comprehension tests. He gave no further comments on whether

such a test would be a 'reading comprehension' test or a measure of those other traits.

Have you ever used cloze tests (blanked tests) for testing comprehension of your students of a passage? How often do you use them? For what other purposes do you use them? What kinds of cloze tests do you use? What do you try to test when you use cloze tests for reading comprehension? What do you think cloze tests measure?

All teachers without exception agreed that they had previously used cloze tests to measure reading comprehension. While some teachers (university and secondary school teachers) had designed cloze tests themselves, others (those teaching in private language centres) had used ready-made cloze tests. The frequency with which teachers had used cloze tests for measuring comprehension differed among teachers. While some had used them as frequently as "*anytime I have taught reading*", "*always*", "*whenever possible*", "*for each new lesson*," and "*in any exam*", others did not seem to have used it so frequently: "*not often*", "*every final or mid-term*", "*not very often*", and "*twice each semester*".

Except for testing reading comprehension, they also reported having used cloze tests for measuring vocabulary and grammar; language knowledge; teaching; word power and power of thinking; measuring proficiency in reading; measuring background knowledge; and as a learning tool for grammar. Three teachers mentioned that they used cloze tests only for measuring reading comprehension and for no other purposes. Most of the cloze tests teachers used were 'standard' cloze tests. A good number of teachers used m/c 'standard' cloze tests, and a few 'rational' cloze tests in which the focus was on particular structures or vocabulary items.

When they used cloze for testing reading comprehension, teachers tried to measure students' understanding of different meaning types (detailed, general, and inferential); "*general and integrated proficiency of subjects*"; "*utilization of discourse-level constraints and structural*

constraints within sentences"; general comprehension; vocabulary meaning; structural knowledge; details; grammar and meaning; and general vocabulary. Evidently, while teachers asserted that they used cloze tests for measuring comprehension, at least half of them had been trying to measure something different from reading comprehension.

Teachers believed that, in general, cloze tests measured the following abilities: language skills such as vocabulary, grammar and reading; general language knowledge; reading comprehension; "*some kind of knowledge (mostly topical and common sense)*"; word usage; range of English knowledge; different things depending on cloze type; literal and general meaning; meaning, grammar, and vocabulary; "*everything*"; background knowledge; and memory. One teacher felt that he did not know what cloze tests measured. Quite strangely, while only 3 believed that cloze tests could measure reading comprehension only, all of them had used cloze tests for this purpose, a purpose for which the majority did not believe they were appropriate. Whether the reason for such widespread use of cloze as a measure of EFL reading comprehension in Iran's secondary schools, and language centres is based on an 'informed' choice or a 'blind practice' is not known from the data in this research, however.

Which of the above methods that you use for testing reading comprehension do you think can measure the comprehension of a text properly? Why?

Generally, most teachers felt that no single technique should be regarded as the best technique because each one had its own advantages and disadvantages. A combination of different techniques was thought to yield a better result. Individual teachers seemed to favour specific techniques for various reasons. Here are some of the techniques they liked and the reasons they provided: different techniques because of different objectives, materials taught, and learners' age; random-deletion cloze because it tests general language; m/c questions, because when students read, there is nothing to observe and because they test

understanding the core of a text; cloze tests, because students should understand overall text meaning to choose the best answer; wh-questions, "*because they are extensive and can test all of the components of a reading text,*" and because students have to answer using their own sentences; and wh-questions and summary because they require full understanding. Only a minority of the teachers were happy with cloze tests as appropriate measures of reading comprehension.

Which testing method do you think that your students prefer? Why?

Almost all the teachers reported that their students preferred questions which required less activity on their part such as choosing an option in m/c questions rather than writing a full sentence as an answer to a question or writing a word in a cloze blank. They believed that students favoured easy-to-handle m/c, T/F, and short-answer questions. Two teachers thought that their students liked cloze tests as measures of reading comprehension. One teacher commented that students cared little about tests. Although the majority of student population whose preferences were discussed by teachers referred to non-university students, the interview and questionnaire data we obtained from university students (the results of which are reported somewhere else) indicated that most of the university students did not favour the techniques thought to be preferred by secondary school students.

Do you think that cloze tests can measure your students' reading comprehension properly? Why?

More than 50% of the teachers believed that cloze tests could measure reading comprehension properly. Although this evidence is somehow in line with the data from the first part of one earlier question in which all the teachers reported having used cloze for testing reading comprehension, it does not quite agree with data for the last part of the same question in which many teachers felt that cloze tests could measure things other than reading comprehension.

Teachers who were positive towards cloze as a measure of reading comprehension reasoned that it made students *"more productive in analysing the passages"*; students were *"involved in the process of learning and producing real-language"*; they were suitable for advanced students who could write properly; those students who were good at meaning, obtained good scores on cloze; and cloze tests could measure reading comprehension properly if they were well-prepared.

The opposite group believed that cloze tests could not measure reading comprehension properly because *"it can't test the student's ability from every point of view"*; some questions could be answered by chance in m/c cloze tests; based on teachers' knowledge about students' proficiency, cloze scores could not show their real level of comprehension; cloze tests mainly focused on grammar and vocabulary; and wh-questions were better than cloze tests. One teacher clearly expressed his view in this way: *"experience has shown me that students don't have adequate knowledge of the language and proper education at getting ... the meaning of a text."*

Do you think that the score that a student gets on a cloze test shows his exact degree of comprehension of the passage? Why?

Except for one teacher who believed that because cloze tests were *integrative* tests and therefore one's scores would show one's comprehension of the passages on which they were based, all others were more or less unanimous that such scores were not indicators of one's real degree of text comprehension. The general tendency was that cloze tests may be better measures of reading comprehension for advanced level students but not for the weaker ones. Talking particularly about m/c cloze tests, most of the teachers felt that the majority of the choices were chosen by chance; other factors than comprehension like general English knowledge and background knowledge affected cloze scores; cloze tests may not be standard; and answers, even if correct, were not based on real comprehension. Some teachers truly felt that there may be no single

"correct comprehension" and that different individuals may arrive at different understandings based on many factors. To make scores more indicative of one's real degree of comprehension, some teachers suggested cloze tests to accompany other reading tests. Whether other reading tests show one's true degree of comprehension of a passage is the focus of the next question.

Do you think that the score that a student gets on other reading comprehension tests shows his exact degree of comprehension of the passage? Why?

One teacher thought that other reading tests could show one's real degree of comprehension. Three more agreed that if questions were properly constructed and were mainly wh-questions, it would be possible that the score one got on a reading test would show his/her exact degree of comprehension. All the rest argued that one's score on a reading test did not show his/her true degree of comprehension along the following lines: "Because comprehension is not an all-or-nothing process, it is inexact"; relativity influences the way knowledge is gauged; "we can never be exact about anything"; direct questions may show a rough understanding only; "some questions measure students' intelligence [rather] than comprehension"; answers may be chosen by chance; students extract whole sentences from the passage and insert them in the answer space without making necessary changes in terms of tense, subject-verb agreement, etc., which shows a superficial matching rather than understanding; carelessness may result in a bad grade; students cannot answer inferential questions; no complete test can test reading comprehension exactly; "measuring exact comprehension is lost to some extent, especially in product type of tests"; and non-comprehension related factors are involved like errors in exam papers, physical context in test session, etc. To all this may be added, from the researcher's own teaching experience, the cheating problem. Cheating is a problem especially in final-term exams when a relatively more important decision is to be made on the success of a student based on the score he/she gets

on a course in many of Iran's secondary schools (and other educational institutions). This problem extends well over to the university context and as noticed by the researcher in this research project, despite having been advised that the tests were for research purposes and there was nothing to do with their achievement or success, some students could not help using the wrong habit of cheating they had acquired. While some cheating problems may be prevented if the test proctors are serious enough, I learned in a recent workshop on Testing and Measurement (July, 2010 at Urmia University) how impossible it may be at times to identify the cheaters when a colleague revealed that a few ladies were using hands free mobile communication, and if the proctor is male, the consequences of getting too close to such test-takers are self-evident.

What other methods do you think can be used to test one's reading comprehension of a given passage in a proper way? Why?

While one teacher believed that there was no perfect way to measure reading comprehension, and that each technique had its own advantages and disadvantages, the rest seemed to believe that certain methods could test reading comprehension more properly than others. "Summary" was mentioned as a proper technique by more than 40% of the teachers. The reasons were: it requires the use of all techniques for understanding a passage; and it also improves writing ability. Some teachers were well aware that because of requiring productive ability, summary may not be suitable for exam purposes, where reading comprehension is only one part of a long exam paper. They also felt that summary was not appropriate for low-proficient readers. Oral summary may be an option, which again requires proficiency in spoken language.

The following techniques were considered proper measures of EFL reading comprehension by other teachers: Preparing good questions and making learners familiar with what they are supposed to do; using 'summary' cloze; using information-transfer tests; a combination of questions intended to test detailed and general meaning; "*a passage with familiar words and fair questions*"; relating words to their sentence-long

explanations given in two lists (which seems to be a test of vocabulary rather than reading comprehension); rewriting the passage in one's own words; making questions about the text one reads; and using a variety of testing techniques. Whether a technique which one 'believes' will measure the real degree of comprehension will truly do the job is another thing, however.

How do you make sure that the questions and techniques you use to test your students comprehension of a passage can test their comprehension properly?

When teachers were asked how they made sure that the questions and techniques they used to test reading comprehension tested comprehension properly, most seemed to have uncertain positions. While one teacher felt that one cannot be sure about it, the majority stated that by following certain procedures and testing 'rules' they could be somewhat sure that they had done the job properly. The latter group felt that the expected result would be achieved by: testing main ideas and intended meaning; testing communicative competence; using a variety of questions, checking response, and looking at it from different angles; asking a similar group to sit the test; giving a "complete" test which would cover everything; getting reliable results; analysing test items after administration and deleting mal-functioning items; and checking to find out if average students could answer comprehension questions.

One teacher commented that in Iran's educational system, for testing reading comprehension "*it is a must to give a certain type of text along with some questions following*". He felt that if there was a 'rule' to follow, the rule should be strong enough to ensure that if one adheres to it, the results would be dependable. For him, adhering to such 'rules' set out by the Ministry of Education would guarantee that the questions he used measured reading comprehension properly. The point, however, is that even if the 'rules' recommended by officials in the Ministry of education are based on sound research evidence (which is not usually the case), they are so generally stated that they may be interpreted and applied in

various forms and that different people may make different tests and each be able to justify their choice according to the rule.

5. Discussion

The preceding analysis of the data gathered through a tailor-made questionnaire indicates the status quo of testing reading comprehension in Iran's secondary schools and language centres where teachers participating in this study came from. Although our findings may not be readily generalisable to other parts of the country, little difference is expected particularly at secondary school level where the 'rules' for all the country are the same. This does not imply in any sense that there is a prescribed system of testing across high schools of the country: teachers are more or less free to exercise their personal preferences in testing, particularly during the term. However, based on our own first-hand experience as teachers at such centres at one stage of our teaching profession and based on our informal observations, we are also confident that similar procedures are adopted by the majority of teachers (at least in the region where this research was completed) during the final-term exam sessions. A surface analysis of the test booklets in Iran's market with samples of Final-Term Examination Papers and various other sample test items—which are in most cases used as models for testing by school teachers—is itself a good proof for our claim.

Regarding our findings about institute teachers, which is based on a limited sample of participants, we are unable to offer any true generalisations for the very reason provided above. This finding applies more readily to the context of the Iran Language Institute and the Jahade Daneshgahi, however, as the participants in our research were sampled from the branches of the above institutes in Urmia. The current practice in these institutes is that all their branches in different cities of Iran follow exactly the same syllabi and the same teaching and testing procedures, more so in the case of the Iran Language Institute. Our findings in the cases of both high school and institute testing systems are in support of those provided by Troudi et al. (2009) who found that EFL

teachers' (as testers) decision on what kinds of tests to construct or use highly depends on their employment conditions rather than being informed by the latest scholarship in the field.

As far as testing reading comprehension at the university level is concerned, our study reveals little. Data from the only university teacher in our sample is in no way representative of what other teachers do in terms of testing reading comprehension in other universities of the country or even at Urmia University. We are therefore not making any claims about how reading comprehension is tested at university level, where, we understand, there is more freedom for teachers to teach and test in the way they find more desirable. Although our findings for high school and institute teachers are open to confirmation by further research — and our hope is that little variation would be found for the similarity of the contexts involved— we propose that interested scholars look at the issue at tertiary education level in greater detail as there does seem to be little uniformity in teaching and testing practices in such centres.

Generally speaking, it seems that not enough importance is given to whether certain testing techniques can really measure what they are intended to. Although initially it appears that educational authorities (of the Ministry of Education or language centres in question) are to blame for setting out the 'rules' without careful research and enough knowledge, we feel that teachers are no less responsible in that they do not exercise enough thought in selecting a test or a technique that is the right one. Contending that teaching and testing are inter-twined and that testing is indeed an integral component of teaching, we cannot agree more with Hughes's (1989) nice justification in this regard that teachers should make their own valuable contributions: "they can write better test themselves, and they can put pressure on others, including professional testers and examining boards, to improve *their* tests" (p. 5). Reflecting on our own experience as language teachers, the fact that most teachers in Iran have to teach many extra hours privately or in other educational institutions because of the economic pressure leaves them with no time and patience to reflect on what is or is not the right teaching or testing

technique for an intended purpose. Most of them, as seen in the questionnaire data, follow what they have been instructed by their teachers at universities decades ago or by the 'rules' set out by their authorities. Although we sympathize with such teachers, we strongly feel that teaching, like any other profession, brings its own commitments, and part of that commitment is the appropriate appraisal of what was supposed to have been delivered to clients, which when fulfilled inappropriately, will mean that the whole education process has been a waste of time and energy.

One clear finding from the evidence presented above is that despite the fact that none of the teachers believed that scores on cloze tests were proper indicators of their students' understanding of the passages on which cloze tests were based, and that only a minority believed that cloze tests were good for measuring reading comprehension, all of them actually used cloze tests (mainly in m/c format) for such a purpose. The reason why teachers have continued to use cloze tests for measuring reading comprehension may be because of the 'rules' they have had to adhere to. This issue is well clear in the case of the private language centres in which teachers are not involved in test construction process, and all tests, including cloze tests, are made ready by 'testing' experts in related organisations, which very occasionally, if ever, make revisions to the tests to accommodate new research findings.

Another major finding of the study is that the majority of teachers believed that the scores their students gained on either cloze tests or any other reading comprehension tests could not truly indicate their degree of comprehension. The implications are that, first thing first, neither cloze tests nor other reading comprehension tests are accurate measures of reading comprehension, and that scores are only figures which may not be reliable indicators of abilities. This evidence in itself undermines the validity of all tests of reading comprehension, including cloze tests, as true measures of text comprehension.

An equally important issue about the above finding is that sometimes scores are so important and decisive in the lives of students,

particularly during final exams, that a student may fail or pass a course if his score is half a point less or more than what he has obtained (and in some cases been given). The implication is that if teachers truly believe that scores do not really show one's true ability, they should not be too serious and inflexible if a student fails only by one score or so. To make the score meanings more valid as true indicators of what one is measuring, it is suggested here that one's pass or fail in a course should not be based on a one-shot score on a test; rather, each student should be examined regularly during the course using a variety of techniques which are appropriate for the relevant purpose, and that the final pass or fail score be based on an average score based on all interim-course and final-course tests. Unfortunately, however, in Iran's education system, although students are routinely given formative tests and a major mid-term test, what mainly counts at the end is the score in the final exam in most cases.

Teachers are, therefore, recommended to place less emphasis on a single score in final exams if important decisions are to be made on a student's future. Although the concern remains that 'busy' teachers will find it difficult to base the achievements of their learners on enough samples of their performance, the suggestion is for teachers to make the most informed decisions (driven by research findings) on what is gained and what is lost when not-carefully-developed measures are used for decision-making purposes. As Bachman (1990) rightly emphasises, a cost-benefit analysis may offer a compromise on the right amount of balance to keep between what has to be sacrificed and what has to be achieved. We also propose that, in addition to quantitative indicators of performance, some assessment-related qualitative information be made available for making important decisions on learners.

The observation that most of the teachers felt that their students would like to be tested through mainly closed techniques like m/c and T/F questions rather than through answering questions seems to be in contradiction with what students themselves thought would be better for them (the report of which appears somewhere else). The fact that the

students talked about by teachers in this paper were mainly secondary school students, who are as a rule less motivated to learn and many of whom attend school because they have to, explains the contradiction. It is, therefore, no surprise that while the majority of secondary school students are expected to favour easy-to-handle tests, most university students are actually interested in questions which would pose some challenge.

The last point to be reiterated is that the results reported and discussed above are based on the ideas of a limited number of teachers not sampled in any systematic procedure. Although we did our best to include a representative sample of ELT teachers at university, high school and institute levels from West Azarbaijan province (and this is why we initially requested 50 teachers to fill out the questionnaires), as an inherent problem with survey-based research (Mackey and Gass, 2005), fewer than expected questionnaires were returned for analysis. The fact that this study was itself a part of larger study made it inconvenient to seek for more participants. Despite the fact that this convenience and opportunistic sampling is widely used in our field for practical considerations, we feel that this is a limitation of our research project. Furthermore, as explained above, although our hunch is that, in the case of high school and institute data, similar observations are predictable for other parts of the country, our strong belief is that with a more rigorous design and a more representative stratified sample (which includes teachers from different regions of the country as well as from various institutes and universities), a more accurate picture of the reality can be depicted. Accordingly, we do not make any claims here to extend our results neither to university contexts, nor to language centres other than the Iran Language Institute and the Jahade Daneshgahi, whose representatives participated in the study.

Also worthy of noting is that since there was no already-validated questionnaire available for such a purpose, the one developed by the researchers may not have been the most appropriate, although it was piloted and found to be acceptable for our purposes as stated earlier. All

these shortcomings mean that the results may not be easily generalisable to other similar contexts across the country, that all the findings and discussions should be interpreted as such bearing in mind the context in which the data for this research was elicited, and that future researchers should take these limitations into consideration in their research. Our research may accordingly be regarded as a starting point for interested researchers to investigate teachers'/learners' beliefs on various aspects of teaching, leaning and testing using the questionnaire introduced here as well as triangulating methodologically through other data elicitation tools including an interview.

6. Conclusion

In light of the evidence provided above regarding teacher judgements as to whether cloze-tests and other measures of reading comprehension employed by teachers at various levels of English education are appropriate for measuring EFL reading comprehension, it can be concluded that neither cloze tests nor other varieties of reading tests investigated here are accurate indicators of text comprehension. Considering the fact that such tests and primarily cloze type tests are relatively widely used not only in Iran but also in well-known international tests such as CAE (Certificate in Advanced English) for the inappropriate purpose of testing reading comprehension, it is suggested that both the 'ruling' bodies and teachers/testers themselves reconsider the use of various forms of cloze tests as measures of reading comprehension particularly when they are used alone. The findings in this piece of research clearly indicate that different cloze tests vary in the degree they can tap overall reading comprehension, and that most of them seem to be testing other constructs than comprehension. Therefore, favouring cloze tests over more valid procedures such as restating the content of the passage or information transfer tasks —simply because the former are more practical and economical than the latter and because cloze results are highly correlated with those of other supposedly valid reading

comprehension measures— is neither substantiated by the evidence in this research nor recommended here.

References

- Abrams, L. M., Pedulla, J.J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of state-wide testing programs. *Theory into Practice, 42*, 18-29.
- Al-Momani, I. A., Ihmeideh, F. M., & Momani, M. (2008). Teachers' views on the effectiveness of United Arab Emirates kindergarten curriculum, instructional strategies, and assessment procedures. *Journal of Research in Childhood Education, 23*, 239-252.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Beycioglu, K., Ozer, N., Ugurlu, C. T. (2010). Teachers' views on educational research. *Teaching and Teacher Education, 26*, 1088-1093.
- Buck, S., Ritter, G. W., Jensen, N. C., & Rose, C. P. (2010). Teachers say the most interesting things – An alternative view of testing. *Phi Delta Kappan, 91*, 50-54.
- Chikasanda, V. K., Ortel-Cass, K., & Jones, A. (in press). Teachers' views about technical education: Implications for reform towards a broad based technology curriculum in Malawi. *International Journal of Technology and Design Education*.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Eret, E., & Ok, A. (in press). Prospective English teachers' views on the physical environment, human resources, and program of their departments. *Asia Pacific Education Review*.
- Gialamas, V., & Nikolopoulou, K. (2010). In-service and pre-service early childhood teachers, views and intentions about ICT use in early childhood settings: A comparative study. *Computers and Education, 55*, 333-341.

- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Levin, T., & Nevo, Y. (2009). Exploring teachers' views on learning and teaching in the context of trans-disciplinary curriculum. *Journal of Curriculum Studies, 41*, 439-465.
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, New Jersey: LEA.
- Mitchell, M. (1994). The views of students and teachers on the use of portfolios as a learning and assessment tool in midwifery education. *Nurse Education Today, 14*, 38-43.
- Ranta, E. (2010). English in the real world vs. English at school: Finnish English teachers' and students' views. *International Journal of Applied Linguistics, 20*, 157-177.
- Sadeghi, K. (2008). Measuring reading comprehension: The judgemental validity of cloze procedure. *Iranian Journal of Applied Linguistics, 11*, 115-132.
- Stevenson, H. (2009). Teachers' beliefs and view of the curriculum. *International Journal of Learning, 16*, 251-262.
- Troudi, S., Coombe, C., & Al-Hamly, M. (2009). EFL teachers' views of English language assessment in the United Arab Emirates and Kuwait. *TESOL Quarterly, 43*, 546-555.
- Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal, 36*, 299-316.
- Yilmaz, K. (2009). Elementary school teachers' views about the critical pedagogy. *The Asia Pacific Education Researcher, 18*, 139-149.
- Wang, J. R., Kao, H. L., & Lin, S. W. (2010). Preservice teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. *Teaching and Teacher Education, 26*, 522-529.

Appendix

Name (optional):..... Sex:

Age:

School/University/Institution you are teaching in:

Level of the students you are teaching:

Number of years teaching:

University Degree you hold:

1. What do you think 'reading comprehension' is?
2. What is the indication that a student has comprehended a text?
3. What are the factors that you think affect one's reading comprehension?
 - Which ones are the most important for you?
4. What do you think are the similarities between real-life and academic reading comprehension?
5. What do you think are the differences between real-life and academic reading comprehension?
 - Which one (real-life or academic reading) do you consider as a proper instance of reading comprehension? Why?
6. How do you think one's real-life comprehension of a passage is tested?
7. How do you think one's academic reading comprehension is tested?
 - Which one do you consider a proper measure of comprehension of a given text? Why?
8. What methods do you use for testing reading comprehension of your students?
9. What do you try to test when you are testing reading comprehension? (for example, literal meaning, general meaning, details, inferential meaning, etc.)
10. Have you ever used cloze tests (blanked tests) for testing comprehension of your students' of a passage?
 - How often do you use them?
 - For what other purposes do you use them?

- *What kinds of cloze tests do you use?*
- *What do you try to test when you use cloze tests for reading comprehension?*
- *What do you think cloze tests measure?*
- 11. *Which of the above methods that you use for testing reading comprehension do you think can measure the comprehension of a text properly? Why?*
- 12. *Which testing method do you think that your students prefer? Why?*
- 13. *Do you think that cloze tests can measure your students reading comprehension properly? Why?*
- 14. *Do you think that the score that a student gets on a cloze test shows his exact degree of comprehension of the passage? Why?*
- 15. *Do you think that the score that a student gets on other reading comprehension tests shows his exact degree of comprehension of the passage? Why?*
- 16. *What other methods do you think can be used to test one's reading comprehension of a given passage in a proper way? Why?*
- 17. *What are your criteria for choosing a passage when you want to test your students reading comprehension?*
 - *How do you decide on the techniques you use for testing comprehension of that passage?*
 - *How do you decide on the number of questions to include in the tests?*
 - *How do you decide on the type of questions, i.e., those for testing general comprehension, for testing details, etc.?*
 - *How do you make sure that the questions and techniques you use to test your students comprehension of a passage can test their comprehension properly?*

Thanks very much for your assistance.