

Teachers' Conceptions of English Language Assessment in Iranian Junior High Schools

¹Masoud Khalili Sabet*

²Mohammadreza Pourgholamali

Research Paper

IJEAP-2305-1959

Received: 2023-05-03

Accepted: 2023-06-15

Published: 2023-06-18

Abstract: Since the 2012 Iranian Education Reform, major changes have been applied to the English assessment system and methods. After a decade, the efficacy and validity of these revisions and modifications are yet unclear. In the Iranian education system, teachers are the main agents of designing and conducting English assessments. It is believed that analyzing their beliefs and attitudes will provide valuable information regarding the current language assessment status. Therefore, the present study was concerned with investigating the Iranian EFL junior high school teachers' conceptions of English assessment. It also attempted to explore the relationship between teachers' experiences and their conceptions of assessment. To this end, the data were collected using the abridged version of the Teachers' Conception of Assessment (TCoA) Inventory. Ninety-six English teachers of junior high schools in Iran were selected through convenience sampling to answer the TCoA inventory. The results of the descriptive statistics indicated that "Improvement" (M=4.19) and "School Accountability" (M=4.05) were endorsed more than other purposes of assessment. Furthermore, the Kruskal-Wallis test showed that the EFL teachers' conceptions differed by their level of teaching experience only in the case of the "control" construct. The low experience group of English teachers had the least mean score for the controlling aspect of assessment. The results of the study were used to suggest practical implications to educational policymakers to improve the English assessment in the Iranian junior high schools.

Keywords: English Language Assessment, Junior high school, Teachers' Conceptions

Introduction

In the recent years, researchers have become increasingly interested in examining the teachers' conceptions of assessment (Brown & Remesal, 2017; Yates & Johnston, 2018; Chen & Teo, 2020, etc.). Conception refers to people's opinions, beliefs, and attitudes concerning the nature of a particular phenomenon (Thompson, 1992). Since teachers implement or prepare students in a test-driven society, it is impossible to examine assessment issues without taking their roles, responsibilities, and conceptions of assessment into account. Teachers' conceptions of the way student's learning is evaluated and the purpose this evaluation is conducted surely have a significant impact on how they teach and what students learn (Fives & Buehl, 2012). Generally, teachers' conceptions of assessment have been studied to identify different purposes of assessment (Newton, 2007). In a broad sense, the purposes of assessment can be classified as improvement and accountability (Brown & Gao 2015). The purpose of improvement is to uncover weaknesses to make changes, but the goal of accountability is to make a judgment on performance (Brown & Gao 2015). Various studies of teachers' conceptions

¹Associate Professor in Applied Linguistics (Corresponding Author), khalilisabet@guilan.ac.ir; Department of English Language, University of Guilan, Guilan, Iran.

²MA in Teaching English, pourgholamali.mr@gmail.com; Department of English Language, University of Guilan, Guilan, Iran.

focused on studying the tension between different assessments purposes (Brown et al., 2009; Gebril & Brown, 2014, etc.).

Since the Iranian school education reform in 2012, fundamental changes have been applied to the English syllabus and textbooks of junior high schools. Communicative Language Teaching (CLT) has been appointed as the main principle governing English materials by policymakers (Foroozandeh & Forouzani, 2015). These changes also affected the methods and process of English assessments. Students' entire performance is now assessed using the results of formative and summative assessments, which include both oral and written exams (Foroozandeh & Forouzani, 2015). Education policymakers have recently placed a strong emphasis on evaluating communicative competence in all separate speaking, listening, reading, and writing skill areas. This movement seems to be going in the right direction however relevant continuous research is needed to investigate its practical effectiveness (Razavipour & Rezagah, 2018).

It is generally accepted that the new assessment reform has caused several challenges and difficulties for teachers and other stakeholders (Akbari, 2015; Razavipour & Rezagah, 2018, etc.). Also, despite evidence of the negative washback effect, summative evaluations and discrete-point examinations, which assess bits of language knowledge, have dominated English assessments in Iranian schools for decades (Riazi & Razavipour, 2011). To this date, only a limited set of general guidelines are provided for students' assessment (Atai & Mazlum, 2013) and no benchmarks are available for the oral exams (Foroozandeh & Forouzani, 2015). Moreover, material developers believe that such guidelines are limited and sometimes fail to capture the intended assessment policies (Atai & Mazlum, 2013). Therefore, it is yet unclear whether the currently conducted assessments in schools are aligned with the purposes of the course book. English teachers are the main agents of educational policy reform and are at the center of language teaching and assessment in Iranian schools. All English assessments given to students are developed, administered; and corrected by the teachers. It is believed that analyzing teachers' conceptions of the purpose of assessment will reveal the nature and purpose of the language assessment and reflects the priorities and insufficiencies of the assessment reform. Also, it would help teachers, school leaders, and policymakers to implement and develop more suitable methods of assessment (Brown & Gao 2015).

Few studies have investigated the impact and challenges of this new assessment reform (Akbari, 2015; Razavipour & Rezagah, 2018). Furthermore, to the best of our knowledge, no study has specifically examined Iranian junior high school EFL teachers' conception of assessment after the curriculum revision and policy reforming. Therefore, the underlying purpose of this research is to investigate the EFL teachers' conceptions of assessment in Iranian junior high schools in order to find out the purposes of English assessment in these contexts. The relationship between teachers' experience and teachers' conceptions of assessment is also explored to find out if there are any differences between low, mid, or high experienced teachers' TCoA. To this end, the present study seeks to answer the following research questions:

Research Question One: What are the English teachers' conceptions of junior high school assessment?

Research Question Two: Are there any significant differences among different conceptions of assessment with respect to teachers' experience?

Literature Review

Teachers' conceptions regarding the purpose behind the assessment process are of particular interest among researchers (Pishghadam & Shayesteh, 2012). It is believed that teachers' conceptions have a significant impact on how they instruct, what students learn, and how learning is assessed (Thompson, 1992). There is strong evidence that beliefs and conceptions have a greater impact on teaching than experiences and socioeconomic environment. Therefore, teachers' conceptions require explicit

attention as they play an important role in implementing educational policies and reforms in schools. According to Thompson (1992), conception can be defined as the implicit knowledge, such as beliefs, values, preferences, and attitudes of a person toward the nature and purpose of a particular phenomenon. Teachers perceive, react to, and engage with their educational environment with conceptions as a framework (Marton, 1981). Conceptions develop through time as a result of a person's interactions with a phenomenon, and they become the mechanism that shapes a person's reactions to that phenomenon (Fives & Buehl, 2012). Therefore, as teachers' conceptions are generated through everyday experience and they reflect back on it in the form of decisions and behaviors, their conceptions are closely tied to practice (Remesal, 2011). Studies claim that assessment often has a detrimental impact on teaching and learning, and examining instructors' conceptions may be a key factor in identifying and dealing with those undesirable outcomes (Brown, 2005). Additionally, studies have indicated that attempts to increase teachers' assessment literacy and create assessment-related policies are more successful when the nature and structure of teachers' conceptions are made clear (Borko et al., 1997). Xu and Brown (2016) defined teachers' conception of assessment as the belief systems that they have toward the method of examining, testing, or evaluating students learning and the purpose of it. Teachers' conceptions of assessment do not follow a consistent pattern, they are inter-connected, and teachers hold multiple and occasionally contradicting conceptions of assessment (Brown et al., 2009). Nevertheless, many researchers focused on identifying various assessment purposes by evaluating the teachers' conceptions of assessment (Delanshere & Jones, 1999; Newton, 2007; Brown, 2002).

Brown (2002) devised a comprehensive framework that categorized teachers' conceptions about the purposes of assessments under four constructs. The purposes are 1. Improvement, assessment as improvement of teaching and learning. 2. School Accountability, assessment as the process of making schools and instructors accountable for their performance. 3. Student Accountability, assessment as the process of holding learners accountable for their learning. (Brown, 2002). 4. Irrelevant, assessment is essentially irrelevant to the lives and work of teachers and students (Brown, 2002). Three of these conceptions namely, improvement, school accountability, and student accountability are categorized as the major purposes of assessment and the term "anti-purpose" is used for the fourth one (Irrelevant). In a broad sense, by combining the school and student accountability functions, the purposes of assessment can be classified as improvement and accountability (Brown & Gao 2015). Brown et al. (2011) and Chinese teachers identified additional purpose for assessment while working in Chinese contexts where high-stakes public examinations predominate. In this study, the use of assessment to control teachers' pedagogical and curriculum practices was one of the more significant purposes. Brown et al. (2015) utilized the control conception for analyzing Indian schools and two items were included to complement the control items concerning covering the examination prescription.

Brown and his associates have developed several self-report instruments based on the Brown's (2002) TCoA framework. These instruments were employed in several countries and languages such as Hong Kong (Brown et al., 2009), Egypt (Gebriel & Brown, 2013), and Turkey (Vardar, 2010). Various studies have also examined the TCoA toward the secondary school's assessment (Brown, 2005; Remesal, 2011; Segers & Tillema, 2011) and English assessment (Pishghadam & Shayesteh, 2012; Hidri, 2016) by using the TCoA inventory (Brown 2006). Since Pishghadam and Shayesteh (2012), and Hidri's (2016) studies are closely related to the context of the present study, they will be reviewed briefly.

Pishghadam and Shayesteh (2012) explored the Iranian EFL teachers' conceptual assessment beliefs with regard to their educational degree, major, age, gender, and experience. The participants of the study were 103 Iranian EFL instructors from different private language institutions in Mashhad, Iran, ranging in age from 22 to 52. The TCoA inventory (Brown, 2002) was employed for this study, and the data from it had an overall reliability estimate of 0.80 according to Cronbach's Alpha. The mean scores for the 'Student accountability' was higher than the other constructs ($M=4.05$). The mean score for the other conceptions were as follows: 'Improvement' ($M=3.77$), 'Irrelevant' ($M=3.46$), and 'School accountability' ($M=3.42$). According to the findings, teachers thought students are in charge of their own language learning. Furthermore, the teachers' experience ranged between 1 to 20 years

(Mean= 5.5) and teachers with more experience thought that assessment reflected the performance of language institutes (School accountability). In Tunisia, Hidri (2016) explored the EFL secondary and university teachers' conception of assessment using TCoA inventory (Brown, 2006). 336 secondary school teachers and 206 university professors participated in the study. Nearly 50% of school teachers had between 11 to 15 years of teaching experience, while 30% of university professors had between 1 to 5 years of experience. The study found a significant relationship between Accountability and Improvement. The research also offered suggestions for exploring TCoA in an EFL environment.

The broad use of the TCoA III abridged version confirms the validity, reliability, and effectiveness of these inventories for examining teachers' assessment conceptions and attitudes. The results of these studies were variant depending on the context, e.g., "Improvement" was the most highly approved factor in all of the "western" samples, but Iranian teachers favored "Student Accountability" more than "Improvement" and Turkish teachers endorsed "Student Accountability" equally to "Improvement". Therefore, the mentioned self-report survey seems to be suitable for studying Iranian junior high school TCoA toward English assessment.

Methodology

Participants

The participants of this study included a group of male and female teachers who teach English at different junior high schools of Guilan in Iran. The participants (N= 96) have been selected through convenience sampling and were all required to respond to the self-report instrument. The sample was taken from the junior high schools that the researcher has professional relationships with educational authorities. Regarding the educational level of the participants, 56 teachers had Bachelor's degree and 40 teachers had Master's degree. The participants were 61 female (63.3%) and 35 male (36.46%) junior high school teachers who taught English as a foreign language. Most of the teachers (N= 47) were between 45 to 54 years old. Thirty-one of the participants had an age range of 34 to 44 and 18 of them were between 23 to 33 years old. The majority of the participants of the study (29.17%) had between 16 to 20 years of teaching English experience as a foreign language. Compared with other categories, there were fewer less experienced (11–15 years).

Questionnaire

The Teacher Conceptions of Assessment version III abridged (Brown, 2006) inventory was employed to gather the quantitative data required for the study. Additionally, another conception of assessment, control, was added from Brown et al. (2011) study. Therefore, the final TCoA scale that is utilized in the present study includes 32 items, 10 factors, and five inter-correlated conceptions (purposes) of assessment: improvement, student accountability, school accountability, irrelevance; and control.

This instrument uses a positively packed agreement rating scale which consists of two negative options such as mostly and strongly disagree and four positive agreement degrees such as slightly, moderately, mostly, and strongly agree. Brown (2006) claimed that the abridged version of TCoA provides similar quality information as the full-scale instrument (TCoA III) and that it is more efficient. This instrument was validated with a national sample of New Zealand secondary school teachers (Brown, 2011). This survey seems to be suitable for examining Iranian junior high school teachers' conception of assessment in an EFL context as it has been formerly employed by Hidri (2016) to examine EFL secondary teachers' TCoA and by Pishghadam and Shayesteh (2012, 2014) to investigate Iranian EFL teachers' TCoA. The results of these studies show that the TCoA inventory is applicable in EFL and Iranian context (Hidri, 2016; Pishghadam & Shayesteh, 2012).

Procedure

First of all, the survey was translated to Persian and a group of academic advisors (four associate professors of university of Guilan with specialization in English language teaching) have been asked to confer its validity. The necessary modifications have been applied and a pre-test of the self-reported survey was carried out with 25 participants of the target population. The alpha-Cronbach reliability of the question was calculated. The estimated values of Cronbach's alpha for the five subscales of the questionnaire came to (α school accountability= .81; α student accountability= .88; α Improvement= .76; α Irrelevance = .89; and α control= .75). The values of Cronbach's alpha revealed that the reliability values of the "school accountability", "student accountability", and "Irrelevance" were "very good" and the reliability of the "improvement" and "control" were "acceptable" based on the reliability standards suggested by (DeVellis, 1991). Finally, an online version of the survey was distributed among the intended participants.

Results

The main quantitative data was gathered through administering the attitude questionnaire to 96 EFL teachers that sought their opinion in five aspects (i.e., school accountability, student accountability, improvement, irrelevance, and control). Then, the data were fed into the statistical package for social science (SPSS) Version 26 for the statistical analyses. For the first research question, descriptive statistics including means and standard deviations were computed for the 32 items of the attitude questionnaire to present a pattern of teachers' views toward high school assessment. As concerns the second research question; the inferential statistics, namely, Kruskal-Wallis H was run to examine the possible differences among teachers with different teaching experiences in terms of their conceptions toward the assessment. Descriptive statistics were computed for the five subscales of the attitude questionnaire. The results are given in Table 1.

Table 1

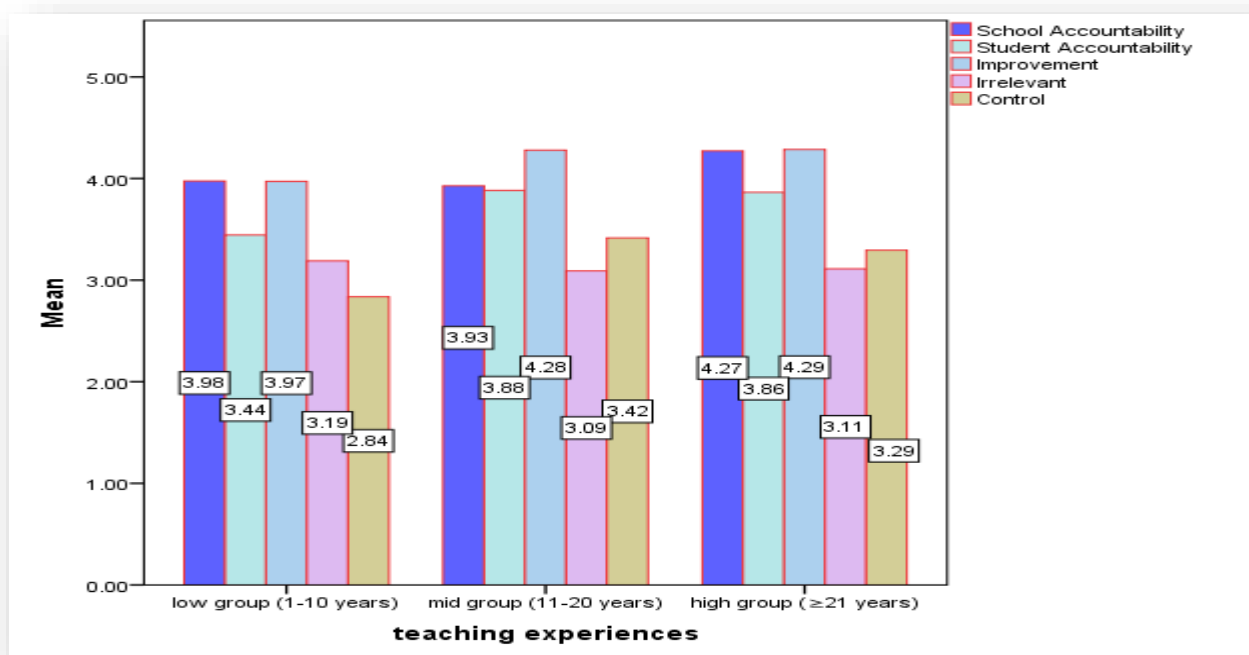
Descriptive Statistics for the Five Subscales of the Attitude Questionnaire

	N	Mean	Std. Deviation	Minimum	Maximum
School Accountability	96	4.0556	.89921	1.67	6.00
Student Accountability	96	3.7535	1.09890	1.00	6.00
Improvement	96	4.1953	.79993	1.17	6.00
Irrelevant	96	3.1250	.65492	1.78	4.67
Control	96	3.2125	1.00139	1.00	5.80

The results indicated that the teachers held more positive attitudes toward the "improvement" (M=4.19; SD= .79) as well as "School Accountability" (M=4.05; SD=.89) than the other aspects of the assessment. Furthermore, the mean score distribution of TCoA across teaching experiences is illustrated in figure 1.

Figure 1

Teachers' Conceptions of High School Assessment across Teaching Experiences



After summarizing the data through the descriptive statistics, Kruskal-Wallis Test was performed on the results of the questionnaires. First, mean ranks were compared for the teachers' conceptions across teaching experiences. The findings are presented in the following section.

Table 2

Ranks for the Teachers' Perceptions of the Assessment

	Experiences	N	Mean Rank
School Accountability	low group (1-10 years)	27	44.48
	mid group (11-20 years)	37	45.32
	high group (≥21 years)	32	55.56
	Total	96	
Student Accountability	low group (1-10 years)	27	39.33
	mid group (11-20 years)	37	52.89
	high group (≥21 years)	32	51.16
	Total	96	
Improvement	low group (1-10 years)	27	38.52
	mid group (11-20 years)	37	53.53
	high group (≥21 years)	32	51.11
	Total	96	
Irrelevant	low group (1-10 years)	27	51.00
	mid group (11-20 years)	37	46.31
	high group (≥21 years)	32	48.92

Control	Total	96	
	low group (1-10 years)	27	37.19
	mid group (11-20 years)	37	53.93
	high group (≥ 21 years)	32	51.77
	Total	96	

The reported mean rank for the three groups revealed that the high group (more than 21 years teaching experience) had the highest overall ranking for “School Accountability” and the low group (1-10 years teaching experience) had the lowest mean rank. In addition, an inspection of the mean ranks for the three groups showed that, with respect to "Student Accountability", the mid group (11 to 20 years teaching experience) had the highest mean rank, with the low group (1-10 years teaching experience) reporting the lowest mean rank. In relation to the “Improvement”, the mid group (11 to 20 years teaching experience) had the highest overall ranking and the low group (1-10 years teaching experience) had the lowest mean rank. With respect to irrelevance, the low and mid groups had the highest and the lowest mean ranks, respectively. And finally as concerns the “Control” category, the mid and low groups had the highest and the lowest mean ranks, respectively. Kruskal Wallis Test was performed to examine whether these rank differences were statistically significant. The results are given in Table 3.

Table 3

Kruskal Wallis for the teachers' Perceptions of the Assessment

	Test Statistics ^{a,b}				
	School Accountability	Student Accountability	Improvement	Irrelevant	Control
Chi-Square	3.151	4.175	4.967	.455	6.342
DF	2	2	2	2	2
Asymp. Sig.	.207	.124	.083	.797	.042
a. Kruskal Wallis Test					
b. Grouping Variable: experience					

The significant results for the teachers' attitudes toward assessment were all higher than .05 except for the last category indicating that there was not a statistically significant difference between EFL teachers in terms of their attitudes towards high school assessment in areas such as “School Accountability, Student Accountability, Improvement, and Irrelevance” across teaching experiences. Nevertheless, for the control category, the sig value was slightly lower than .05 implying that there was a statistically significant difference among teachers concerning their perceptions of high school assessment in the area of “control” across teaching experiences ($p \leq .05$).

Discussion

As it was mentioned before, the Junior high school teachers highly supported ‘Improvement’ as the purpose of English assessments. Also, the ‘Irrelevant’ construct had the least endorsement among other aspects of assessment. This may indicate that generally the English teachers of Iranian junior high schools believe that assessment is relevant and aids in improving both English proficiency and all-around character of the learners. School accountability received the teachers' second-highest endorsement. Given that Iranian schools are in competition with one another and their quality is determined by the students' test performance, this is fairly acceptable in the context of the research. As a result, teachers' beliefs are in line with Iranian educational system since they think that a good school improves students' learning and teaching and this is reflected in students' exam scores.

These findings are in line with previous studies of TCoA in Hong Kong (Brown et al. 2009), Egypt (Gebriel & Brown, 2013), and India (Brown et al., 2015). This similarity between the conceptions of the Iranian junior high school teachers and the previous studies may refer to the high-stakes, centralized education systems of these countries where examinations is used as a method to control

schools, learners, and instructors and rely on public tests to determine admission to higher education. Therefore, the teacher's belief that the tests improve students' learning and motivation is justified in the examination-driven context of the study.

Furthermore, the results of this study are in contrast with the previous TCoA study in Iranian context. In Pishghadam and Shayesteh (2012) study, the highest endorsed conception was the "student accountability" conception whereas the least mean score was attributed to the "school accountability". The differences between the two researches may relate to the context in which the data was gathered. It seems that the private institutes' instructors hold students responsible for their own learning while public school teachers believe that the assessment results show the success or failure of the school's performance. Nevertheless, it appears that rather than highlighting the role of students or schools, teachers should concentrate on improving their instructional strategies by utilizing the outcomes of formative assessments to enhance students' learning. Also, the mean score for "Improvement" is less than the present study. This may indicate that the latest reform of the Iranian system of Education and English instruction of schools in 2012 had some impact on instructors' ability to use assessment to enhance both their instruction and students' learning in the classroom. However, this movement toward improvement and formative assessment is weak. The instructors still need time and direction to alter their views of accountability by placing less of a focus on assessment as an evaluation tool and more of an emphasis on a more diagnostic approach.

The analysis of the sub-factors of the TCoA survey also showed that teachers had positive views regarding the trustworthiness and productiveness of assessment in providing feedback to students about their abilities. However, they also believed that assessment is a good and reliable source to evaluate a school and places students in clusters according to their abilities. In addition, many of the participants accepted that "assessment controls the content of teachers' classes". It seems that in the Iranian education environment, the 'accountability' purpose of assessment is not only an indicator of school, student, or teachers' quality and performances but also to some extent it relates to controlling them.

The participants of this study were grouped by their years of teaching experience in to three groups: low group (10 years of experience or less), mid group (between 10 to 20 years) and high group (20 years or more). The Kruskal-Wallis H test results revealed that the EFL teachers' conceptions differed by their level of teaching experience but these differences were statistically significant simply in terms of "control" feature. This is in contrast with the results of the Pishghadam and Shayesteh (2012) study. Pishghadam and Shayesteh (2012) claimed that there were no significant differences between different TCoAs and teacher experience. Although both studies were conducted in the Iranian EFL context, it appears that the results from the two studies are not fully comparable. Since Pishghadam and Shayesteh (2012) did not include the "control" construct in their data collection process, it is yet unclear whether the private English teachers' conceptions regarding controlling aspects of assessments differ by their teaching experience or not. Further research is needed to analyze private English teachers' beliefs and attitudes toward the controlling features that the private institutes might include in the assessment procedures.

Furthermore, the low experience group of teachers in this paper had the least mean score for the 'Control' construct. It can be assumed that the new generation of English teachers do not accept the controlling aspect of assessment and believe that the summative function of English proficiency assessment should either be replaced with other type of assessment such as the formative ones or be removed completely. This is also evident in low endorsement of 'Improvement' and high support of 'Irrelevance' among the low experienced teachers compared to others. This lack of belief in irrelevance and improving function of assessment is probably originated from poor quality of assessment courses in teacher training program and universities. Finally, the high experienced teachers were more likely to agree to 'School accountability' construct and believed that assessment's results show the performance of the schools. The reason behind this endorsement could be their previous experiences with formal product-based assessment strategies either as a teacher or a student. This is in line with Vardar (2010) findings as the more experienced teachers were in favor of the Student Accountability purposes of

assessment and the researcher similarly claimed that the reason behind this belief is the participants' previous experiences.

Conclusion and Implications

The first research question intended to identify the Iranian English teachers' conceptions of junior high school assessment. Generally, teachers believe that a good school improves English learning and teaching which is reflected in students' exam scores. As Akbari (2015) pointed out, since school authorities and parents think that a school's education quality is determined by the student's test performance, accountability aims of assessment may dominate teachers' assessment beliefs.

The second research question dealt with the relationship between teachers' experience and their conceptions of assessment. The EFL teachers' conceptions differed by their level of teaching experience only in the case of the "control" construct. Therefore, the null hypothesis was partially rejected and there was a statistically significant attitudinal difference among teachers with different teaching experiences. Apparently, attitudes toward the accountability and control purposes of assessment are less evident in the low experienced groups of teachers than more experienced ones. Instead, this new generation of English instructors believes that the English assessment is irrelevant, impeding the improvement of learner's language learning and skills.

To conclude, the teachers believe that most of the challenges of conducting English assessment can be resolved if the policymakers pay attention to the assessment of English as a foreign language and the beliefs toward traditional summative language assessments are altered among teachers, students, parents, and other stakeholders

Various implications can be drawn from the teachers' conceptions for policymakers to improve the English assessment in the Iranian junior high schools and to solve the current assessment related challenges. First and foremost, the teacher training programs, such as seminars or workshops should be improved in a way that shift teachers' conceptions from accountability and control purposes of assessment to improvement and more meaningful ones. Furthermore, since the new generation of teachers lacks the belief in importance and efficacy of English assessment, it seems that the university testing classes should be analyzed and modified to teach pre-service teachers the function and methods of evaluating learners' language skills.

Further research is required to shed light on the conceptions and beliefs toward English assessment in the Iranian context. The participants in the present study were selected from junior high school teachers of Guilan in Iran through convenience sampling. To ensure a more reliable generalization of TCoA, it is recommended that the participants be selected from all over the country's schools, and- that stratified random sampling be used. Also, other English instructors, as well as teachers' beliefs and conceptions toward assessment at the private language institutes, universities should be analyzed and compared with each other to achieve a broader picture of Iranian TCoA.

Additionally, other stakeholders, including students, parents, school authorities, and educational policy makers' conceptions of assessment can be researched. Furthermore, female and male conceptions can be compared to find out if there is a difference between them. Finally, the effect of different beliefs and conceptions toward assessment on teaching methods, classroom practices, and students' achievements can be studied.

Acknowledgement

We would like to thank all those who helped us conduct the research.

Declaration of Conflicting Interests

The authors do not have any conflicts of interest to declare.

Funding Details

The authors did not receive any funding any financial support from any organization or agency.

References

- Akbari, Z. (2015). Current challenges in teaching/learning English for EFL learners: The case of junior high school and high school. *Procedia: Social and Behavioral Sciences*, 199, 394-401.
- Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching and Teacher Education*, 13(3), 259-278.
- Brown, G. T. L. (2002). *Teachers' conceptions of assessment*. Unpublished doctoral dissertation, University of Auckland, Auckland
- Brown, G. T. L. (2005). An exploration of secondary school teachers' conceptions of assessment. In B. Kozuh, T. Beran, A. Kozłowska & P. Bayliss (Eds.), *Measurement and assessment in educational and social research* (pp. 39-48).
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged version. *Psychological Reports*, 99, 166-170.
- Brown, G. T. L. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters*, 3, 45-70.
- Brown, G. T. L., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, 2(1)
- Brown, G. T. L., & Remesal, A. (2017). Teachers' conceptions of assessment: Comparing two inventories with Ecuadorian teachers. *Studies in Educational Evaluation*, 55, 68-74.
- Brown, G. T. L., Chaudhry, H., & Dhamija, R. (2015). The impact of an assessment policy upon teachers' self-reported assessment beliefs and practices: A quasi-experimental study of Indian teachers in private schools. *International Journal of Educational Research*, 71, 50-64.
- Brown, G. T. L., Hui, S. K., Flora, W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307-320.
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347-363.
- Chen, J., & Teo, T. (2020). Chinese school teachers' conceptions of high-stakes and low-stakes assessments: An invariance analysis. *Educational Studies*, 46(4), 458-475.
- Delandshere, G., & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: A case of assessment paralysis. *Journal of Curriculum and Supervision*, 14(3), 216.
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.

- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the “messy” construct of teachers’ beliefs: What are they? Which have been examined? What can they tell us? *In American Psychological Association eBooks* (pp. 471–499). <https://doi.org/10.1037/13274-019>
- Foroozandeh, E., & Forouzani, M. (2015). Developing school English materials for the new Iranian educational system. In Kennedy, C. (Ed.), *English language teaching in the Islamic Republic of Iran: Innovations, trends and challenges*, (pp. 59-70). British Council.
- Gebril, A., & Brown, G. T. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian teachers’ conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 21(1), 16-33.
- Hidri, S. (2016). Conceptions of assessment: Investigating what assessment means to secondary and university teachers. *Arab Journal of Applied Linguistics*, 1(1), 19-43.
- Marton, F. (1981). Phenomenography—Describing conceptions of the world around us. *Instructional Science*, 10(2), 177-200.
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, 14(2), 149-170.
- Pishghadam, R., & Shayesteh, S. (2012). Conceptions of assessment among Iranian EFL teachers. *The Iranian EFL Journal*, 8, 9–23.
- Razavipour, K., & Rezagah, K. (2018). Language assessment in the new English curriculum in Iran: Managerial, institutional, and professional barriers. *Language Testing in Asia*, 8(1), 1-18.
- Remesal, A. (2011). Primary and secondary teachers’ conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482.
- Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37(1), 49-54.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 127–146). Macmillan Publishing Co, Inc.
- Vardar, E. (2010). *Sixth, seventh and eighth grade teachers' conception of assessment* (Master's thesis, Middle East Technical University).
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162.
- Yates, A., & Johnston, M. (2018). The impact of school-based assessment for qualifications on teachers’ conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 25(6), 638-654.