Assessment Literacy Components Predicting EFL Teachers' Job Demand-Resources: A Focus on Burnout and Engagement

Much has been done on assessment literacy (AL) components and job demand-resources (JD-R). However, an interdisciplinary look at AL components as the predictors of JD-R and its consequences in the engagement and burnout of teachers' assessment performance has been neglected. To fill this gap, this study explored this issue in the Iranian context. To this end, through convenience sampling, 146 Iranian EFL teachers were selected to answer questionnaires on AL, JD-R, burnout and engagement. A series of multiple regression analyses were run to analyze the collected data. The results showed that some components of AL including test construction, administering, rating, and interpreting test, psychometric properties of a test, using and interpreting statistics and authenticity were significant predictors of job demand. Moreover, the results revealed that alternative and digital-based assessment, recognizing test type, distinction and function, and authenticity were significant predictors of job resource. Furthermore, test construction, administering, rating, and interpreting test, psychometric properties of a test, and using and interpreting statistics could significantly predict teachers' burnout. In addition, alternative and digital-based assessment, giving feedback in assessment, and ethical and cultural considerations in assessment could significantly predict teachers' engagement. These findings can have theoretical and practical implications for stakeholders.

Keywords: *Assessment literacy, Job demand, Job resource, Burnout, Engagement*

Introduction

As assessment literacy is a prerequisite of a successful academic career for an EFL teacher, one should consider that work-related factors could also play a role. Every EFL teacher should be equipped with AL since a considerable amount of teachers' preparation time is allocated to creating instruments and observing learners, marking, recording, and analyzing and synthesizing results in reports in their daily teaching career. The assessment practices used by teachers at schools have been the concern of various studies (e.g., Rogers, 1991; Wilson, 1998, 2000). Some studies have focused on identifying assessment practices in EFL/ESL contexts (e.g., Cheng, et al., 2004). For instance, Fischer (2002) believes that environmental and organizational factors can affect the assessment performance of teachers; yet such factors are often understudied. It has also been observed that assessment performance has been ignored in light of teachers' assessment knowledge/literacy (Chung, 2008).

To better understand organizational factors that relate to assessment performance, L2 teachers' job resources and job demands must be dealt with first. Bakker and Demerouti (2007) believe that job demands have multiple

UNDER PEER REVIEW

aspects. In other words, job demands are nothing but physical, emotional, and cognitive requirements imposed by specific professions (Fernet et al., 2015). In the case of L2 teachers' assessment, it can be defined as how assessment tasks and duties put physical, emotional, or cognitive burdens on the shoulders of teachers.

Job resources are those factors that either lessen job demands or reduce the psychological or physiological costs associated with each job or those aspects that have a functional role in achieving goals in working performance or the factors that may motivate learning and personal development (Schaufeli & Bakker, 2004). Although Hobfoll (2002) points out that the role of resources is not limited necessarily to job demands, those factors have an important position in their own right.

In line with the above explanations, if an assessment task is physically, emotionally, and cognitively daunting, it may affect the assessment performance of language teachers. Furthermore, if an assessment task was done in the situation satisfactorily, it would influence teachers' assessment performance and lead to engagement. Such an interdisciplinary look at the assessment performance of assessors has been neglected in the assessment literature. In this study, the researchers extended the JD-R model, proposed by Demerouti et al. (2001), by measuring engagement and burnout independently as possible consequences of assessment tasks either in the form of job resources or job demands in teachers' assessment performance. The following research questions were raised and explored:

1. Which of the components of assessment literacy can better predict EFL teachers' job demands?
2. Which of the components of AL can better predict EFL teachers' job resources?
3. Which of the components of AL can better predict EFL teachers' burnout?
4. Which of the components of AL can better predict EFL teachers' work-engagement?

Literature Review

Assessment Literacy

Assessment literacy is among the most significant predictors of teachers' development in their teaching profession. According to Webb (2002), AL refers to language teachers' knowledge in assessing language and understanding the results by interpreting them. One may think of Language assessment literacy (LAL) as a repertoire of competencies that make it possible for a teacher to understand, check and, under some circumstances, develop language tests and interpret test results (Pill & Harding, 2013). A number of studies have explored AL and its components. Moreover, the prominence of

assessment in both pre-service and in-service programmes for teacher education has continued to manifest itself in assessment literature.

Brindley (2001) was among the earliest to recognize the importance of AL for language testers and educators, and the issue emerged as a theme at conferences and other fora (Hasselgreen et al., 2003; Huhta & Tarnanen, 2007). The studies unanimously showed an increased need for teacher training in more learner-oriented assessment practices. Volante and Fazio (2007) advocated a systematic analysis of potential discrepancies between student-teachers´ assessment curriculum and their actual achievements.

Davies (2008), proposed an extensive set of goals for AL education comprising domains of skills, knowledge, and principles. Following Brindley's model, Inbar-Lourie (2008) outlined a tripartite model aligning practical and theoretical knowledge with a socio-historical understanding of the implications of assessment. Inbar-Lourie, (2008) also underlined the value of intertwining assessment and learning by adhering to assessment-for-learning practices (Black & Wiliam, 1998) and to dynamic assessment (Poehner & Lantolf, 2005) in the field of AL. Popham (2009) attempted to set standards for assessment, particularly in the field of teacher education. He provided a collection of teacher- and student-oriented statements incorporating practical assessment knowledge and skills.

Voss et al. (2011) tested an overall framework for student teachers' psychological/pedagogical knowledge and suggested empirical structures entailing the following kinds of knowledge relating to: classroom management, classroom assessment, teaching methods, the learning processes, and learner characteristics. At the same time, Brookhart (2011) offered an operationalization of teacher assessment abilities and skills applicable to both classroom assessment and test administration. Giving feedback, communication and scaffolding student autonomy in assessment are informed by Brookhart's work.

Fulcher (2012), incorporated practical knowledge, theoretical knowledge, and socio-historical understanding of assessment-related activities. Fulcher also appreciated student-teachers and their experiences of courses and academic study. Also, he acknowledged that competence at all levels should not be required form all stakeholders. Taylor (2013) suggested different profiles for various groups of stakeholders comprising the following components: technical skills, language pedagogy, knowledge of theory, personal beliefs, principles and concepts, local practices, and decision-making. In a qualitative study, Baker and Riches (2013) looked at the way LAL is developed in teachers and language assessment experts while they were collaboratively engaged in revising national examinations in the Haitian EFL context. The finding revealed that the LAL of both groups developed.

Jeong (2013) explored the way language assessment courses (LACs) were offered in different countries and how the course instructor influenced such courses. The findings showed that the content of LACs may significantly vary depending on factors like the language testing background of the instructors in such areas as test specifications and theories, basic statistics, development of

ONLY FOR REVIEW

rubric, classroom assessment, and accommodation of test. Lam (2015) attempted to shed more light on how language assessment courses are offered in Hong Kong and the way such courses influence pre-service teachers' development of LAL. The findings showed that training in language assessment remained inadequate and the gap between theory and practice within the context of assessment reform was not bridged even by selected LACs.

Xu and Brown (2016) revisited knowledge aspects of previous AL models in a large-scale study. The knowledge constituents incorporated included knowledge of assessment purposes, disciplinary knowledge and pedagogical content knowledge, knowledge of grading, content and methods, feedback knowledge, knowledge of interpreting assessment and communication, knowledge of self and peer-assessment, and knowledge of ethics in assessment. Deneen and Brown (2016) examined the effect of an educational assessment course on MA level student teachers and reported that the AL of both pre-service and practicing teachers increased. In another study, Fernando (2018) investigated the effect of literacy assessment in a formative academic context on engaging students in their writing processes through a web learning platform. The findings indicated that the use of online technology to conduct literacy assessment in a formative academic context can reduce students' fear of written assessment and improve their writing.

The developments in the area of AL in recent years, and the increased expectations of learners for explicit assessment, have left teachers with no choice but to become assessment literate. In this regard, Medland (2015) clearly acknowledges the non-mature state of AL, which needs to be accompanied by more literature in the higher education context. William (2015) echoes a similar sentiment. Similar to other contexts, the Iranian EFL context has a long path to reach the ideal status in the AL of language teachers (Mellati & Khademi, 2018).

Job Demand-Resource and its Consequences

Much has been done on job demand-resource and its consequences either as engagement or burnout. For example, Jansen et al. (2020) investigated how motivation and psychological need satisfaction can account for the way job resources and job demands affect teachers' learning commitment. The findings showed positive relationships between satisfaction of basic psychological needs, teachers' job resources experience, commitment to professional learning, and autonomous motivation. However, there was no relationship between basic need satisfaction job and demands. In another study, De Carlo et al. (2019) examined the how work-related factors (in the form of job resources and job demands) were associated with work-family conflict (WFC) in teachers. They reported that workload (in both forms) had a positive relationship with WFC, and job resources buffered this association.

Additionally, Yin et al. (2018) investigated the factors predicting teachers' well-being at both individual and school levels by incorporating individual

factors into the JD-R model. The findings suggested that emotional job demands of teaching at the school-level and suppression at the individual level has positive relationships with teachers' depression and anxiety, while trust in colleagues (school-level) and reappraisal (individual-level) had positive relationships with contentment and enthusiasm. Emotional job demands also turned out to be positively associated with suppression. Also, Saleem et al. (2017) studied the correlations between principals' leadership styles and teachers' organizational commitment and reported a significant relationship between them.

Furthermore, Khan et al. (2014) reviewed the associations between job resources, job demands, and burnout empirical studies through a non-systematic existing literature review. They recommended that factors contributing to JD-R and burnout be located and controlled in an organized way. Also, Ismail et al. (2009) studied the relationship between job performance and occupational stress, focusing on emotional intelligence among academicians. They found that job performance was related to occupational stress and that emotional intelligence mediated this relationship.

Winefield et al. (2003) studied occupational stress among Australian university staff. They found that general staff were better off in comparison with academic staff. They also found that the new staff experienced more strain and less job satisfaction. In addition, self-report measures of psychological wellbeing were correlated with objective measures of university well-being. Also, Aimi Roslan et al. (2015) investigated teachers' work engagement and burnout. The results indicated that job demands and job resources were negatively correlated. The results also suggested a positive relationship between burnout and job demands.

Shaikh et al. (2018) analyzed the effect of personal resources and job resources on university teachers' job engagement. They found that both factors had a notable role in predicting work engagement. Also, Jagodics and Szabó (2014) found a positive relationship between burnout and job demands. However, job resources and burnout were negatively correlated. They further reported that burnout was negatively related to professional and emotional support of co-workers. The results imply that certain factors of the workplace contribute to the development of burnout, whereas others appear to lessen the impact of job demands. Furthermore, the findings indicated that burnout was significantly affected by the social environment of the workplace.

Javed and Cheema (2015) found that organizational resources such as marketing capability, technology and financial resources increase work engagement. In addition, they observed that work engagement and organizational resources increase service climate, which in turn, improves the performance of employees.

Vera et al. (2012) analyzed the role of self-efficacy in predicting the JD-R Model. The researchers longitudinally examined the two underlying processes including the motivational process and the erosion process among teachers. The findings confirmed both processes and the predictive power of self-efficacy.

ONLY FOR REVIEW

Salmela-Aro and Upadyaya (2018) examined the relationships among personal and JD-R, engagement and work burnout during career stages. They also assessed the relationship between general well-being with burnout and work engagement. The findings showed that, particularly during the early stages of career, economic problems appeared to be linked to burnout symptoms. However, in the later stages, caregiving demands had a positive relationship with work burnout but a negative association with work engagement. The findings further suggested that, during early career stages, ICT demands were positively related to work burnout. Also, life satisfaction was related to work engagement while depressive symptoms were linked to work burnout in all career stages.

Grayson and Alvarez (2008) evaluated the effect of components of school climate on the dimensions of emotional exhaustion, including depersonalization and feelings of low personal accomplishment. Results indicated that each burnout subscale was related to different school climate-related criteria. Similarly, Prieto et al. (2008) attempted to predict burnout among teachers based on the JD-R model by adding personal resources. The results showed that quantitative overload has a predictive role in exhaustion and dedication.

Bayani et al. (2003) investigated Iranian high school teachers' sex, age, and years of experience to determine which factor was more likely to lead to burn-out. The results indicated that male teachers were more likely to be infected with burnout compared to female teachers. This study emphasized the significance of designing a well-structured and evaluative burnout reduction program. Also, Gonzalez-Roma et al. (2006) examined the relationship between work engagement and burnout to see if they were independent factors. The results showed that work engagement and burnout were, in fact, opposite factors. In a similar study, Schaufeli et al. (2008) reported a positive correlation between burnout and workaholism. However, there was no significant relationship between workaholism and work engagement. In another study, Skaalvik (2020) studied the perceptions of school principals' regarding job resources and demands. The results showed seven potential moderately-correlated job resources and nine moderately-correlated job demands factors. Moreover, it was reported that, of the potential resources and demands, only four were significantly related to emotional exhaustion, job satisfaction, and motivation to quit. Meanwhile, emotional exhaustion and job satisfaction mediated the relationships between motivation to quit and JD-Rs.

From the above review, it can be noted that many studies have looked at teacher assessment through giving prominence to different components and combinations of them. Meanwhile, several studies have considered the effects of JD-R and its consequences in the form of burnout and engagement in different professions. However, few studies, if any, have linked JD-R to AL, especially in the form of assessment performance. This study is undertaken with the aim of addressing this gap.

ONLY FOR REVIEW

Methodology

Participants and Setting

To collect data, 146 Iranian EFL teachers (67 males and 79 females) from different provinces of Iran including Markazi, Qazvin, Tehran, Fars, Khuzestan, and Mazandaran were selected through convenience sampling. The participants were English language teachers with teaching experience ranging from four to 40 years. Their age range was between 21 and 65 years. The sample consisted of 22 B.A. holders, 75 M.A. students or MA holder and 49 PhD students or PhD holders.

Instrumentation

The following data collection instruments were utilized to meet the objectives of this study.

The Researcher-made Assessment Literacy Questionnaire

The researcher-made questionnaire included 35 items assessing 9 components of AL such as: test construction (5 items), administering, rating and interpreting test (5 items), alternative and digital-based assessment (4 items), giving feedback in assessment (3items), ethical and cultural considerations (5 items), psychometric properties of a test (3 items), using and interpreting statistics (4 items), recognizing test types, distinction and function (3 items), and authenticity of a test (3 items). This 5-point Likert questionnaire required teachers to show their knowledge in each of these items. The researcher estimated the reliability of the newly developed questionnaire using Cronbach's alpha (.87). Moreover, the content validity of the instrument was checked through expert judgment.

Burnout Questionnaire

This study utilized 'Copenhagen Burnout Inventory; CBI' to assess academic burnout (Kristensen, et al., 2005). In this study, only the personal burnout scale was used. This scale refers to the degree of exhaustion (psychological and physical) experienced by a person (Kristensen, et al., 2005). This scale contains six questions. Language teachers were given a five-point Likert scale to indicate their answer, ranging from 1 (never) to 5 (always). Its validity was checked through expert judgment, and its reliability was estimated using Cronbach alpha to be .83.

Engagement Questionnaire

Language teachers' job engagement was assessed using the short version of the Utrecht Work Engagement Scale (UWES) (Schaufeli & Salanova, 2007).

ONLY FOR REVIEW

It had three subscales including 17 statements that reflected the underlying dimensions of engagement: vigour (6 items), dedication (5 items), and absorption (6 items). All items were scored on a 7-point frequency rating scale ranging from 0 (never) to 6 (always). Schaufeli and Bakker (2004) reported correlations among all dimensions from 0.90 to 0.95. The researcher estimated the reliability of this scale using Cronbach alpha (0.87).

Job Demand and Resources Questionnaire

The job demand questionnaire includes two main sections; the first one (job demand) includes seven parts, each of which includes four questions: pace and amount of work, mental load, emotional load, physical effort, changes in tasks, ambiguities about work, and uncertainty about future. The second one (job resource) also involves seven parts, each of which includes four questions: participation, information, communication, relationship with superior, relationship with colleagues, remuneration, and independence in the work. Both parts are 5-point Likert scale ranging from 1(never) to 5 (always). The researchers checked its validity through expert judgment and estimated its reliability (0.83), using Cronbach alpha.

Procedure and Data Analysis

The researchers extended the JD-R model which was proposed by Demerouti et al. (2001), by measuring engagement and burnout independently as possible consequences of assessment tasks either in the form of job demands or job resources in teachers' assessment performance. Initially, the researchers developed an AL questionnaire that assessed various aspects of AL and distributed it among 386 Iranian EFL teachers. A Principal Component Analysis (PCA) was carried out, and a questionnaire with 35 items which evaluated nine components of ALwas developed.

Then, the validated version of the researcher-made questionnaire and the JD-R, burnout and engagement questionnaires were administered to 220 EFL teachers via e-mail, social messaging applications like WhatsApp, Telegram, and Google forms. The process of distributing the questionnaires and the method of collection was as follows. First, the participants' s consent was sought, and the researchers explained that the process of data collection is such that each respondent must answer four questionnaires with time intervals to both prevent fatigue and help them achieve the desired goal. The participants were given opportunities to ask questions if they had any. Then, the four questionnaires were sent out to the teachers. Despite completing the consent form, only 146 of the participants filled out all the four questionnaires. To answer the research questions, a series of multiple regression analyses were conducted.

ONLY FOR REVIEW

Results

The First Research Question

The first question aimed at examining the predictive power of the components of AL over job demand scores. To answer this question, a multiple regression analysis was used. The first step was checking its assumptions. The Durbin-Watson statistic was used to check the independence of residuals. Table 1 shows the result.

Table 1. Durbin-Watson statistic for Checking Independence of Observations for Job Demand

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	1.000[a]	.999	.999	.284	2.053

In this table, the value of Durbin-Watson statistic is 2.053, which is in the range of 1.5 to 2.5. Therefore, it can be concluded that the data are independently observed. The second assumption is multicollinearity. To test this assumption, the correlation between each pair of AL components was checked, the results of which are presented in Table 2.

Table 2. Analysis of Multicollinearity between Each Pair of Independent Variables

	F1	F2	F3	F4	F5	F6	F7	F8	F9
F1	1	-.102	-.017	.448*	-.417*	-.084	.447*	-.124	-.046
F2			.573*	-.174*	-.018	.551*	-.060	.060	-.129
F3			1	-.178*	-.139	.599*	.050	.085	-.040
F4				1	-.217**	-.073	.271*	-.044	.055
F5					1	-.009	-.404*	.065	-.027
F6						1	.056	.104	.041
F7							1	-.019	.097
F8								1	-.070
F9									1

As the coefficients of correlation in Table 2 indicate, there is no high correlation between any pairs of components; thus, the assumption has been met. The third assumption is homoscedasticity, which was checked by examining the scatterplot; there was no sign of the violation of this assumption. After checking the assumptions, the researchers ran multiple regression using the standard Method. Model Summary is presented in Table 3.

ONLY FOR REVIEW

Table 3. Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	1.000a	.999	.999	.284

a. Predictors: (Constant), F9, F5, F6, F8, F4, F7, F2, F1, F3

Table 4 reports the ANOVA which assesses the overall significance of the multiple regression model. It shows that the model is significant ($F_{(9, 136)}$ = 20854, P < .05).

Table 4. ANOVA for Significance of the Multiple Regression Model of AL Components and Job Demand

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15124.034	9	1680.448	20854.248	.000
	Residual	10.959	136	.081		
	Total	15134.993	145			

Table 5 shows the standardised coefficients. It indicates that the p-values for five components of AL, including test construction, administering, rating, and interpreting test, psychometric properties of a test, using and interpreting statistics and authenticity are less than 0.05. This means that these components of AL are significant predictors of job demand with coefficients of 0.422, 0.443, 0.422, 0.338, and 0.290, respectively.

Table 5. Standardised Coefficients for Job Demand

Model		Standardized Coefficients Beta	t	Sig.
1	(Constant)		1.366	.174
	F1	.422	142.997	.000
	F2	.443	145.852	.000
	F3	-.002	-.584	.561
	F4	-.003	-1.000	.319
	F5	-.004	-1.441	.152
	F6	.422	135.468	.000
	F7	.338	123.963	.000
	F8	-.003	-1.430	.155
	F9	.290	121.931	.000

The Second Research Question

Investigating the predictive power of job resource over the components of AL was the aim of the second question. To answer this question, a multiple regression analysis was run. Prior to that, the assumptions were checked. The

Durbin-Watson statistic confirmed the assumption of independence of residuals.

Table 6. Durbin-Watson statistic for Checking Independence of Observations of AL Components and Job Resource

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	1.000[a]	.999	.999	.300	2.160

As Table 6 indicates, the value of Durbin-Watson statistic is 2.160, which is in the range of 1.5 to 2.5. So, the data are independently observed. Multicollinearity was already checked in Table 2. The assumption of homoscedasticity was also checked, and the scatterplot showed that there was no violation. Then, multiple regression analysis was run using Enter Method to determine the significant predictors of job resource. Table 7 presents the model summary.

Table 7. Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	1.000a	.999	.999	.300

a. Predictors: (Constant), F9, F5, F6, F8, F4, F7, F2, F1, F3

The ANOVA which assessed the overall significance of the multiple regression model showed that the model was significant ($F_{(9, 136)} = 2338$, $P < .05$). Table 8 shows the standardised coefficients. It shows that three components of AL (i.e., alternative and digital-based assessment, recognizing test type, distinction and function, and authenticity) were significant predictors of the participants' job resource scores with coefficients of 0.582, 0.680, and 0.438, respectively.

Table 8. Standardised Coefficients for Job Resource

Model		Standardized Coefficients Beta	t	Sig.
1	(Constant)		1.130	.261
	F1	-.001	-.194	.847
	F2	-.001	-.466	.642
	F3	.582	193.015	.000
	F4	-.003	-1.087	.279
	F5	-.003	-1.056	.293
	F6	.002	.694	.489
	F7	-.003	-1.294	.198
	F8	.680	306.387	.000
	F9	.438	194.686	.000

UNDER PEER REVIEW

The Third Research Question

The third research question investigated the predictability of the burnout scores of the participants by the components of AL. To answer this question, a multiple regression analysis had to be run. Like the previous research questions, the first step was checking its assumptions. The Durbin-Watson statistic was checked for the assumption of independence of residuals. The result (Durbin-Watson statistic = 1.86) showed that the value was in the range of 1.5 to 2.5. The second assumption (multicollinearity) was checked and its results were reported in Table 2. After checking the scatterplot for the assumption of homoscedasticity, multiple regression analysis was done. The model summary is presented in Table 9.

Table 9. *Model Summary for Burnout*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.996a	.992	.992	.279

a. Predictors: (Constant), F9, F5, F6, F8, F4, F7, F2, F1, F3

ANOVA results ($F_{(9,\ 136)}$ = 1941, P < .005) showed that the model was significant Therefore, the results of the standardised coefficients could be considered. Table 10 shows that four components of AL (i.e., test construction, administering, rating, and interpreting test, psychometric properties of a test, and using and interpreting statistics) were significant predictors of burnout with multiple regression coefficients of 0.498, 0.531, 0.509, and 0.024, respectively.

Table 10. Standardised coefficients for Burnout

Model		Standardized Coefficients	t	Sig.
		Beta		
1	(Constant)		-2.085	.039
	F1	.498	51.726	.000
	F2	.531	53.492	.000
	F3	-.004	-.390	.697
	F4	.000	-.044	.965
	F5	.010	1.194	.235
	F6	.509	50.024	.000
	F7	.024	2.733	.007
	F8	.003	.330	.742
	F9	.001	.065	.948

The Fourth Research Question

In order to examine the predictability of the engagement scores of the participants by the components of AL, which was the objective of the fourth

research question, the researchers ran a multiple regression analysis on the data. As usual, prior to running this test, its assumptions were checked. The Durbin-Watson statistic was checked for the assumption of independence of residuals; the result (Durbin-Watson statistic = 2.053) showed that the assumption was met. The multicollinearity assumption was checked in Table 2. The assumption of homoscedasticity was also checked.

As the necessary assumptions for running the multiple regression analysis were met, the researchers ran this test using Enter Method to determine the components of the assessment literacy which predict the work-engagement scores of the participants. Model summary is presented in Table 11.

Table 11. Model Summary for Engagement

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999a	.998	.998	.271

a. Predictors: (Constant), F9, F5, F6, F8, F4, F7, F2, F1, F3

The significance of the regression model was checked, and ANOVA results ($F_{(9, 136)}$ = 9769, P < .005) showed that the model was significant and that we could refer to the results of standardised coefficients (Table 12).

Table 12. Standardised Coefficients for Engagement

Model		Standardized Coefficients Beta	t	Sig.
1	(Constant)		1.600	.112
	F1	-.002	-.407	.685
	F2	.001	.304	.762
	F3	.639	137.042	.000
	F4	.599	153.665	.000
	F5	.662	169.197	.000
	F6	.002	.417	.677
	F7	-.009	-2.234	.027
	F8	.003	.779	.437
	F9	.475	136.644	.000

Table 17 indicates that four components of AL, including Alternative and Digital-based Assessment, Giving Feedback in Assessment, and Ethical and Cultural Considerations in Assessment) were significant predictors of the participants' engagement scores with multiple regression coefficients of 0.639, 0.599, 0.662, and 0.475, respectively.

Discussion

This study revealed that some components of AL including test construction, administering, rating, and interpreting test, psychometric properties of a test, using and interpreting statistics and authenticity were significant predictors of job demand. Interestingly, social, cultural, and psychological factors were not among these factors. This can be associated with the fact that still traditional beliefs are being held by many second language teachers. This finding seems to be in line with those of some previous studies (e.g., Aimi Roslan et al., 2015; Jagodics & Szabó, 2014) This similarity between the findings of this study and those of previous research can be explained with the fact that job demands are nothing but physical, emotional, and cognitive requirements imposed by specific professions (Fernet et al., 2015). If an assessment task is physically, emotionally, and cognitively daunting, it may affect the assessment performance of language teachers. Another reason for the similarity might be that the more quantitative and cognitively demanding the task assessment, the more difficult it would be to handle. As a result, those observed AL components as predictors of job demand need more time and amount of work and higher levels of information processing to handle.

These findings are also in contrast with those of some other studies. For instance, Mulder (2017) has argued that psychological stress is an important factor. Also, Maertz et al. (2007) have stated that social and psychological factors are significantly associated with job demand. This contrast can be linked to the educational context of Iran, where many teachers tend to hold traditional beliefs about language teaching and assessment. Another reason might be the neglect of psychological and sociocultural issues by managers, which need to be addressed.

It should be acknowledged that some previous studies have provided more comprehensive explanations in that they have considered factors such as psychological characteristics (Jansen et al., 2020) as well as exhaustion and motivation (Skaalvik, 2020) as important factors. Hence, we acknowledge this shortcoming of our study. The reason can be associated with the fact that this study focused mainly on the test itself, rather than the participants. Moreover, some studies have focused on factors outside school. For instance, De Carlo et al. (2019) considered work-family conflict as one of the important factors.

Moreover, this study showed that some components of AL were significant predictors of job resource. The factors of AL which predict job resources (alternative and digital-based assessment, recognizing test type, distinction and function, and authenticity) are associated with assessment itself. This shows that Iranian teachers do not seem to pay much attention to the role of sociocultural and psychological factors. This is inconsistent with Skaalvik (2020), who argued that psychological factors are important as well. The inconsistency may be due to the differences between the contexts where the two studies were conducted. It seems that there are some differences between teachers in different contexts. Similar to Slaalvik, De Carlo et al. (2019)

ONLY FOR REVIEW

considered protective factors as an important issue regarding job resources. Also, Yu et al. (2018) reported that psychological factors are of crucial importance. However, the participants of neither this study nor previous studies have considered social factors as an important element regarding job resources.

In some ways, this finding seems to be in line with that of Javed and Cheema (2015), who observed that work engagement and organizational resources increase service climate, thus improving employee performance. The similarity between the findings of this study and those of others can be explained with the fact that job resources constitute those psychological, physical, social, and organizational aspects of the job that are useful in achieving work objectives, lowering job demands, and stimulating personal growth (Bakker & Demerouti, 2007; Schaufeli & Bakker, 2004). Furthermore, Fischer (2002) believes that environmental and organizational factors can affect the assessment performance of teachers. Therefore, it could be concluded that alternative and digitally-based assessment, as one of the technological resources at the service of the teachers, could be seen as a valuable resource to deal with assessment tasks.

This study also showed that we need to pay more attention to factors other than those associated with the test itself – such as those related to the individuals and their families. In this regard, there seem to be some similarities between the findings of the present study and those of previous ones. For instance, Yin et al. (2018) have argued that personal factors are important. Especially, they have focused on anxiety and depression. In similar veins, Shaikh et al. (2018) have argued that personal and psychological factors are crucial. De Carlo et al. (2019) have considered families as well. However, they have simply focused on work-family conflict and have not considered other issues associated with teachers' families. Hence, it can be predicted that in the future, more studies will focus on the sociocultural dimension. This problem exists in the context of Iran as well (Khanjani et al., 2016; Mellati & Khademi, 2018; Razavipour et al., 2011). In the context of Iran, only recent studies have considered either the social or the psychological aspect. For instance, Moradan and Pourasadollah (2014) have focused on the psychological aspect, albeit only partially. Also, Zolfaghari and Ahmadi (2016) have paid a little attention to the social aspect. However, as findings of this study showed, more research should be done in this regard.

Also, this study showed that test construction, administering, rating, and interpreting test, psychometric properties of a test, and using and interpreting statistics can significantly predict teachers' burnout. Similar to the previous questions, the results revealed that the participants considered traditional properties as more important. Previous research suggests that social and psychological factors are the main sources of burnout. For example, Jadgodics and Szabo (2014) reported that emotional issues and peer conflicts are among the most important sources of teachers' burnout. Moreover, they argued that emotional and professional social support of co-workers can prevent burnout. Similarly, Vera et al. (2012) found that emotional and social factors such as autonomy, social support and self-efficacy are important factors when dealing

UNDER PEER REVIEW

with burnout. Similarly, Maslach and Leiter (2016) have argued that burnout is caused by chronic physical, emotional, and mental stresses. Therefore, it seems necessary to inform Iranian EFL teachers about the importance of psychological and social factors.

Like the previous questions, there are contradictions between the findings of this question and those of the previous studies. For instance, Bayani et al. (2013) have introduced personal factors like gender as important. Similarly, engagement was introduced by Gonzalez-Roma et al. (2006) as another important factor to predict burnout. Grayson and Alvarez (2008) considered school environment as an important factor. The reason for the differences may be the fact that while the present study has not paid enough attention to personal and contextual factors in predicting teachers' burnout, previous studies have not paid enough attention to factors related to the test itself.

This study also showed that some factors of AL including alternative and digital-based assessment, giving feedback in assessment, and ethical and cultural considerations in assessment can significantly predict teachers' engagement. Contrary to the previous research questions, the responses to this question revealed participants' attention to cultural and ethical considerations in addition to traditional issues. However, social and emotional issues still tend to be ignored. This is inconsistent with some of the previous studies. For instance, Salmela-Aro and Upadyaya (2018) argued that life satisfaction and depression are associated with burnout. Roslan et al. (2015) reported that job demands such as pupils' misbehavior are a source of teachers' burnout. Also, Javed and Cheema (2015) argued that service climate is an important source of burnout.

Since work engagement and motivation are the two faces of the same coin (Gagné, 2014), because work engagement is actually the fulfilling state of mind that is work-related and is characterized by dedication,vigor, and absorption (Schaufeli et al., 2002), surprisingly, those AL components that predict engagement are those factors which can increase motivation in the relationship between student and teacher in assessment tasks such as giving feedback, or paying attention to cultural issues. Moreover, one of the components that predict engagement (alternative and digital-based assessment) is the predictor of job resource, too. This is consistent with Vera et al. (2012), who reported that job resources foster engagement.

Conclusion and implications

By extending the JDR model to assessment performance, this study indicated that assessment related performance such as test construction, administering, rating, and interpreting test, psychometric properties of a test, using and interpreting statistics and authenticity, being on the shoulder of second language teachers, especially in examination-oriented contexts such as Iran, can lead to job demands. These physical, social, or organizational demands require more emotional, psychological, physical, and mental effort.

ONLY FOR REVIEW

The finding also showed that AL components that predict job resources such as alternative and digitally-based assessment as one of the technological resources at the service of the teachers could be seen as a valuable resource to deal with assessment tasks. Also, this study showed that test construction, administering, rating, and interpreting test, psychometric properties of a test, and using and interpreting statistics can significantly predict teachers' burnout. The reason is that the four components of AL that predict job demands are among the predictor of burnout. Moreover, this is in line with the idea that in the job stress context, the strain like burnout is considered as a sort of a deleterious condition, which is caused by job stressors/demands (Jain et al., 2013).

On the other hand, the findings indicated that some factors of AL including alternative and digital-based assessment, giving feedback in assessment, and ethical and cultural considerations in assessment can significantly predict teachers' engagement. Those AL components that predict engagement are those factors which can increase motivation in the relationship between student and teacher in assessment tasks such as giving feedback, or paying attention to cultural issues. Moreover, one of the components that predict engagement (alternative and digital-based assessment) is the predictor of job resource, too. This shows that job resources significantly influence teachers' engagement.

All in all, it appears that in the context of Iran, not only are teachers affected by job demands such as test construction, administering, rating, and interpreting test, psychometric properties of a test, using and interpreting statistics and authenticity in their assessment performance - but also they are not provided with enough job resources in assessment performance such as alternative and digitally-based assessment as one of the technological resources. Furthermore, factors such as alternative and digital-based assessment, giving feedback in assessment, and ethical and cultural considerations in assessment as predictors of engagement in assessment performance should be included as essential components of successful assessment performance.

The finding of this study can be used by authorities in order to design teacher assessment training programs which attempt to provide teachers with enough assessment resources to reduce the burden of assessment tasks (demands) on the shoulders of teachers. Teachers themselves also can use the finding of this study to become more familiar with assessment resource aspects that will lead to motivation in their assessment task and prevent them from burnout. Moreover, teachers can try to design and interpret authentic tests, digital-based assessment, and computer-based assessment, each of which offers an opportunity to access job resources and get away from burnout.

ONLY FOR REVIEW

Acknowledgments

The current authors would like to thank all participants who contributed to the data collection, as well as the anonymous reviewers and experts who helped us to improve this work.

References

Aimi Roslan, N., Ann Ho, J., Imm Ng, S., & Sambasivan, M. (2015). Job demands and job resources: Predicting burnout and work engagement among teachers. *International Proceedings of Economics Development and Research*, *84*, 138-145. URL: http://www.ipedr.com/list-111-1.html

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309-328. https://doi.org/10.1108/02683940710733115

Bayani, A. A., Bagheri, H., & Bayani, A. (2013). Influence of gender, age, and years of teaching experience on burnout. *Annals of Biological Research*, *4*(4), 239-243.

Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honor of Alan Davies* (pp. 126–136). Cambridge University Press.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, *30*(1), 3–12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Cheng, L., Watanabe, Y., & Curtis, A. (2004). Washback in language testing: Research contexts and methods. *Applied Linguistics*, *26*(3), 459-462. https://doi.org/10.1093/applin/ami015

Chung, R. R. (2008). Beyond assessment: Performance assessments in teacher education. *Teacher Education Quarterly*, *35*(1), 7-28. http://www.jstor.org/stable/23479028

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, *25*(3), 327-347. https://doi.org/10.1177%2F0265532208090156

De Carlo, A., Girardi, D., Falco, A., Dal Corso, L., & Di Sipio, A. (2019). When does work interfere with teachers' private life? An application of the job demands-resources model. *Frontiers in Psychology*, *10*, 1-13. https://psycnet.apa.org/doi/10.3389/fpsyg.2019.01121

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands resources model of burnout. *Journal of Applied Psychology*, *86*(3), 499–512. https://psycnet.apa.org/doi/10.1037/0021-9010.86.3.499

Deneen, C. C., & Brown, G. T. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, *3*(1), 1-14. http://dx.doi.org/10.1080/2331186X.2016.1225380

Fernando, W. (2018). Show me your true colors: Scaffolding formative academic literacy assessment through an online learning platform. *Assessing Writing*, *36*, 63–76. https://doi.org/10.1016/j.asw.2018.03.005

Fernet, C., Trépanier, S.G., Austin, S., Gagné, M., & Forest, J. (2015). Transformational leadership and optimal functioning at work: On the mediating role of employees' perceived job characteristics and motivation. *Work & Stress*, *29*(1), 11-31. https://doi.org/10.1080/02678373.2014.1003998

UNDER PEER REVIEW

Fischer, T. B. (2002). Strategic environmental assessment performance criteria—The same requirements for every assessment? *Journal of Environmental Assessment Policy and Management*, *4*(1), 83-99. https://doi.org/10.1142/S1464333202000905

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, *9*(2), 113-132. https://doi.org/10.1080/15434303.2011.642041

Gagné, M. (2014). *The Oxford handbook of work engagement, motivation, and self-determination theory*. Oxford University Press. https://doi.org/1093/oxfordhb/9780199794911.001.0001

Gonzalez-Roma, V., Schaufeli, W. B., Bakker, A. B., & Lloret, S. (2006). Burnout and work engagement: Independent factors or opposite poles? *Journal of Vocational Behavior*, *68*(1), 165-174. https://doi.org/10.1016/j.jvb.2005.01.003.

Grayson, J. L., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education*, *24*(5), 1349-1363. https://doi.org/10.1016/J.TATE.2007.06.005

Hobfoll, S. E. (2002). Social and psychological resources and adaptation. *Review of General Psychology, 6*(4), 307–324. https://doi.org/10.1037%2F1089-2680.6.4.307

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, *25*(3), 385–402. https://doi.org/10.1177%2F0265532208090158

Ismail, A., Suh-Suh, Y., Ajis, M. N. E., & Dollah, N. F. (2009). Relationship between occupational Stress, emotional intelligence and job performance: An empirical study in Malaysia. *Theoretical & Applied Economics*, *16*(10), 3-16.

Jagodics, B., & Szabó, E. (2014). Job demands versus resources: Workplace factors related to teacher's burnout. *Practice and Theory in Systems of Education, 9*(4), 1-15.

Jain, A. K., Giga, S. I., & Cooper, C. (2013). Perceived organizational support as a moderator in the relationship between organizational stressors and organizational citizenship behaviors. *International Journal of Organizational Analysis*, *21*(3), 313-334. http://dx.doi.org/10.1108/IJOA-Mar-2012-0574

Jansen in de Wal, J., van den Beemt, A., Martens, R.L., & Den Brok, P. J. (2020). The relationship between job demands, job resources and teachers' professional learning: Is it explained by self-determination theory? *Studies in Continuing Education, 42* (1), 17–39. https://doi.org/10.1080/0158037X.2018.1520697

Javed, F., & Cheema, S. (2015). The relationship between organizational resources and work engagement: The mediating role of service climate as a predictor of performance and loyalty in shopping malls of Pakistan. *Journal of Business Studies Quarterly, 7*(2), 11-22.

Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing, 30*(3), 345–362. https://doi.org/10.1177%2F0265532213480334

Khan, F., Yusoff, R.M.D., & Khan, A. (2014). Job demands, burnout and resources in teaching a conceptual review. *World Applied Sciences Journal, 30*(1), 20-28. http://dx.doi.org/10.5829/idosi.wasj.2014.30.01.21029

Khanjani, A., Vahdany, F., & Jafarigohar, M. (2017). EFL teacher education in Iran: Does it promote trainees' pedagogical content knowledge? *Journal of Research in Applied Linguistics, 8*(2), 159-186. https://dx.doi.org/10.22055/rals.2017.13096

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. http://dx.doi.org/10.1177/0265532214554321

Maertz, C. P., Griffeth, R. W., Campbell, N., & Allen, D.G. (2007). The effects of perceived organizational support and perceived supervisor support on employee turnover. *Journal of Organizational Behavior*, *28*(8), 1059–1075. http://dx.doi.org/10.1002/job.472

Maslach, C., & Leiter, M. P. (2016). Understanding the burnout experience: Recent research and its implications for psychiatry. *World Psychiatry*, *15*(2), 103-111. https://dx.doi.org/10.1002%2Fwps.20311

Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education*, *43*(6), 1-18. http://dx.doi.org/10.14221/ajte.2018v43n6.1

Moradan, A., & Pourasadollah, R. (2014). Teachers' thinking about their teaching: A critical study on Iranian TEFL teachers. *Procedia - Social and Behavioral Sciences, 98*, 1194–1203. http://dx.doi.org/10.1016/j.sbspro.2014.03.534

Mulder, M. (2017). Competence theory and research: A synthesis. In M. Mulder (Ed.), *Competence-based vocational and professional education. Bridging the worlds of work and education* (pp. 1071-1106). Springer.

Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing, 30*(3), 381–402. https://doi.org/10.1177%2F0265532213480337

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice, 48*(1), 4–11. https://doi.org/10.1080/00405840802577536

Prieto, L.L., Salanova, M., Martínez, I.M., & Schaufeli, W.B. (2008). Extension of the job demands-resources model in the prediction of burnout and engagement among teachers over time. *Psicothema*, *20*(3), 354-360.

Razavipour, K., Riazi, A., & Rashidi, N. (2011). On the interaction of test washback and teacher assessment literacy: The case of Iranian EFL secondary school teachers. *English Language Teaching, 4*(1), 156-161. https://doi.org/10.5539/elt.v4n1p156

Saleem, Z., Batool, S., & Khattak, S. R. (2017). Relationship between leadership styles and organizational commitment: Moderating role of emotional intelligence and organizational support. *Journal of Managerial Sciences*, *11*(1),71-84.

Salmela-Aro, K., & Upadyaya, K. (2018). Role of demands-resources in work engagement and burnout in different career stages. *Journal of Vocational Behavior, 108*, 190–200. https://doi.org/10.1016/j.jvb.2018.08.002

Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior, 25*(3), 293-315. https://doi.org/10.1002/job.248

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, *3*(1), 71-92. https://doi.org/10.1023/A:1015630930326

Schaufeli, W. B., Taris, T. W., & Rhenen, W. V. (2008). Workaholism, burnout, and work engagement: Three of a kind or three different kinds of employee well-being? *Applied Psychology*, *57*(2), 173-203. https://doi.org/10.1111/j.1464-0597.2007.00285.x

Shaikh, S.S., Talpur, U., Muhammad Khuwaja, F., & Katper, N.K. (2018). Assessing the effect of personal and job resources on work engagement of university teachers.

International Journal of Advanced and Applied Sciences, 5(11), 100-109. https://doi.org/10.21833/ijaas.2018.11.015

Skaalvik, C. (2020). Self-efficacy for instructional leadership: Relations with perceived job demands and job resources, emotional exhaustion, job satisfaction, and motivation to quit. *Social Psychological Education*, 23, 1343–1366. https://doi.org/10.1007/s11218-020-09585-9

Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing, 30*(3), 403–412. http://dx.doi.org/10.1177/0265532213480338

Vera, M., Salanova, M., & Lorente, L. (2012). The predicting role of self-efficacy in the job demands-resources model: A longitudinal study. *Estudios de Psicología*, *33*(2), 167-178. https://psycnet.apa.org/doi/10.1174/021093912800676439

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, *30*(3), 749–770. https://doi.org/10.2307/20466661

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, *103*(4), 952–969. https://psycnet.apa.org/doi/10.1037/a0025125

Winefield, A. H., Gillespie, N., Stough, C., Dua, J., Hapuarachchi, J., & Boyd, C. (2003). Occupational stress in Australian university staff: Results from a national survey. *International Journal of Stress Management*, *10*(1), 51-63. https://psycnet.apa.org/doi/10.1037/1072-5245.10.1.51

Xu, Y., & Brown, G. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teacher and Teacher Education*, *58*, 149-162. http://dx.doi.org/10.1016/j.tate.2016.05.010

Yin, H., Huang, S., & Lv, L. (2018). A multilevel analysis of job characteristics, emotion regulation, and teacher well-being: A job demands-resources model. *Frontiers in Psychology*, *9*, 23-95. https://psycnet.apa.org/doi/10.3389/fpsyg.2018.02395

Zolfaghari, F., & Ahmadi, A. (2016). Assessment literacy components across subject matters. *Cogent Education*, *3*(1), 1-16. http://dx.doi.org/10.1080/2331186X.2016.1252561

ONLY FOR REVIEW