ORIGINAL ARTICLE

The Washback Effects of Task-based and Non-Task-based Language Assessment of English for Academic Purposes (EAP) on Students' Reading Comprehension Ability

Maryam Nazari[1], Abbas Bayat[2,*], Peyman Rajabi[2]
[1]Ph.D. Candidate, English Department, College of Humanities, Malayer Branch, Islamic Azad University, Malayer, Iran
[2]English Department, College of Humanities, Malayer Branch, Islamic Azad University, Malayer, Iran

*Malayer Branch, Islamic Azad University, Daneshgah Blvd. Malayer, 65718/117 Iran

Tel: +98 9183521796
Email: abbasbayat305@gmail.com

Background: The present study set out to examine the washback effects of task-based (TBLA) and non-task-based language assessment (NTBLA) on nursing students' reading comprehension ability.

Method: First, 60 students studying nursing in Kermanshah University of Medical Sciences took part in the study. Then, they were homogenized in terms of English language proficiency and were later non-randomly assigned to two experimental groups each comprising 30 learners. In Group1, task-based reading assessment was provided while in Group 2, non-task-based traditional reading assessment was administered.

Results: The results of the study pointed to a statistically significant difference between the washback effects of the two assessment types. In other words, the participants in TBLA group outperformed those in NTBLA group.

Conclusion: Complementing EAP courses for medical students with TBLA sounds an auspicious move towards enhancing academic reading ability of medical students. The findings have certain pedagogical and research implications as well.

Keywords: English for Academic Purposes, Students of Medicine, Task-based Language Assessment, Non-task-based Language Assessment, Reading Ability

التأثيرات العكسية لتقييم اللغة المبني على المهام وغير المبني على المهام في اللغة الإنجليزية للأغراض الأكاديمية (EAP) على قدرة الطلاب على الفهم القرائي

الخلفية: تهدف الدراسة الحالية إلى فحص تأثيرات الارتداد لتقييم اللغة القائم على المهام (TBLA) وتقييم اللغة غير القائم على المهام (NTBLA) على قدرة استيعاب القراءة لدى طلاب التمريض.

الطرق: أولاً، شارك في الدراسة 60 طالباً يدرسون التمريض في جامعة كرمانشاه للعلوم الطبية. ثانياً، تم تجانسهم من حيث إتقان اللغة الإنجليزية وتم توزيعهم لاحقاً بشكل غير عشوائي على مجموعتين تجريبيتين تضم كل منهما 30 متعلماً. في المجموعة 1، تم توفير تقييم القراءة على أساس المهام بينما في المجموعة 2، تم إجراء تقييم القراءة التقليدية غير القائمة على المهام.

النتائج: أشارت نتائج الدراسة إلى وجود فرق ذو دلالة إحصائية بين تأثيرات الاسترجاع لنوعي التقييم. بمعنى آخر، تفوق المشاركون في مجموعة TBLA على المشاركين في مجموعة NTBLA.

الاستنتاج: إن استكمال دورات EAP لطلاب الطب مع TBLA يبدو خطوة ميمونة نحو تعزيز قدرة القراءة الأكاديمية لطلاب الطب. النتائج لها بعض الآثار التربوية والبحثية أيضا.

الكلمات المفتاحية: اللغة الإنجليزية للأغراض الأكاديمية، طلاب الطب، تقييم اللغة على أساس المهام، تقييم اللغة غير القائم على المهام، القدرة على القراءة

بررسی اثرات ثانویه ارزیابی زبان تکلیف محور (TBLA) و غیر تکلیف محور (NTBLA) بر توانایی درک مطلب دانشجویان

زمینه و هدف: این مقاله به بررسی اثرات ثانویه ارزیابی زبان تکلیف محور (TBLA) و غیر تکلیف محور (NTBLA) بر توانایی درک مطلب دانشجویان پرستاری می پردازد.

روش: ابتدا 60 نفر از دانشجویان رشته پرستاری دانشگاه علوم پزشکی کرمانشاه برای شرکت در این مطالعه انتخاب شدند. سپس از نظر مهارت زبان انگلیسی همگن شدند و بهطور غیرتصادفی به دو گروه آزمایشی که هر کدام شامل 30 زبان‌آموز بود، تقسیم شدند. در گروه یک، ارزیابی خواندن تکلیف محور ارائه شد، در حالی که در گروه دو، ارزیابی خواندن سنتی غیر تکلیفی انجام گردید.

یافته‌ها: نتایج مطالعه از وجود تفاوت معنی‌دار آماری بین اثرات ثانویه دو نوع ارزیابی حکایت داشت. به عبارت دیگر، شرکت کنندگان در گروه TBLA از گروه NTBLA بهتر عمل کردند.

نتیجه‌گیری: تکمیل دوره‌های انگلیسی برای اهداف دانشگاهی (EAP) برای دانشجویان علوم پزشکی با TBLA حرکتی مبارک در جهت افزایش توانایی خواندن آکادمیک این دانشجویان به نظر می‌رسد. یافته ها دارای مفاهیم آموزشی و پژوهشی خاصی نیز هستند.

واژه های کلیدی: انگلیسی برای مقاصد تحصیلی، دانشجویان پزشکی، ارزیابی زبان تکلیف محور، ارزیابی زبان غیر تکلیف محور، توانایی خواندن

تعلیمی مقاصد کے لیے انگریزی پر ٹاسک بیسڈ اور نان ٹاسک بیسڈ لینگویج اسسمنٹ کے واش بیک اثرات

پس منظر: موجودہ مطالعہ ٹاسک بیسڈ (ٹی بی ایل اے) اور نان ٹاسک بیسڈ لینگویج اسسمنٹ (این ٹی بی ایل اے) کے نرسنگ طلباء کی پڑھنے کی فہم کی صلاحیت پر واش بیک اثرات کا جائزہ لینے کے لیے ترتیب دیا گیا ہے۔

طریقہ کار: پہلے، کرمانشاہ یونیورسٹی آف میڈیکل سائنسز میں نرسنگ کی تعلیم حاصل کرنے والے 60 طلباء نے مطالعہ میں حصہ لیا۔ دوسرا، انہیں انگریزی زبان کی مہارت کے لحاظ سے ہم آہنگ کیا گیا تھا اور بعد میں 30 سیکھنے والوں پر مشتمل دو تجرباتی گروپوں کو غیر تصادفی طور پر تفویض کیا گیا تھا۔ گروپ 1 میں، ٹاسک پر مبنی پڑھنے کی تشخیص فراہم کی گئی تھی جبکہ گروپ 2 میں، غیر ٹاسک پر مبنی روایتی پڑھنے کی تشخیص کا انتظام کیا گیا تھا۔

نتائج: مطالعہ کے نتائج نے دو تشخیصی اقسام کے واش بیک اثرات کے درمیان اعدادوشمار کے لحاظ سے ایک فرق کی طرف اشارہ کیا۔ دوسرے الفاظ میں، ٹی بی ایل اے کے شرکاء نے این ٹی بی ایل اے گروپ کے شرکاء سے بہتر کارکردگی کا مظاہرہ کیا۔

نتیجہ: TBLA کے ساتھ میڈیکل طلباء کے لئے EAP کورسز کی تکمیل میڈیکل طلباء کی تعلیمی پڑھنے کی صلاحیت کو بڑھانے کی طرف ایک اچھا اقدام ہے۔ نتائج کے کچھ تدریسی اور تحقیقی مضمرات بھی ہیں۔

کلیدی الفاظ: تعلیمی مقاصد کے لیے انگریزی، طب کے طلباء، ٹاسک پر مبنی لینگویج اسسمنٹ، نان ٹاسک پر مبنی لینگویج اسسمنٹ، پڑھنے کی اہلیت

INTRODUCTION

In task-based language assessment (TBLA), which draws substantially on the communicative approach, language use is assessed in contexts that appear to be more authentic and multifaceted than in discrete skills assessment. As Mislevy et al. (2002) argued, TBLA typically necessitates the integration of topical and socio-pragmatic knowledge along with knowledge of the structural elements of language (1).

Brindley (1994, p. 74) defined TBLA as "the process of evaluating, in relation to a set of explicitly stated criteria, the quality of the communicative performances elicited from learners as part of goal-directed, meaning-focused language use requiring the integration of skills and knowledge" (2). The kind of interest in TBLA owes largely to issues such as the orientation of task-based assessment towards task-based language teaching, positive 'washback' effects of assessment practices on teaching/learning, and the limitations associated with discrete-skills assessments (DSAs) (3). DSAs are directed towards the assessment of language knowledge, targeting morphosyntactic and comprehension aspects of such knowledge with discrete-point and more decontextualized test items.

In the field of second/foreign language teaching (L2), many studies have been conducted to examine the effect of various instructional approaches on learners' reading comprehension ability (4-6). When it comes to English for Academic Purposes (EAP), the role of reading comprehension ability becomes outstandingly more salient since reading skill constitutes a fairly major part of EAP courses at least in the context of the current study, i.e., Iran. To this end, the present study aimed at examining the comparative effects of task-based and non-task-based language assessment on Iranian medical students' reading comprehension ability. Another impetus behind the conduct of the present study lies in the scarcity of past related studies, to the best knowledge of the researcher, on the washback impacts of TBLA and non-TBLA on medical students' reading comprehension ability.

With the introduction of TBLT approach, there was a zest to design tests which could assess test-takers' language ability as they engaged in performing authentic tasks (7). Accordingly, TBLA was advanced by a number of testing specialists (1, 6, 8, 9). Brindley (1994) refers to TBLA as "the process of evaluating, a set of explicitly stated criteria, and the quality of the communicative performances elicited from learners as part of goal-directed, meaning-focused language use requiring the integration of skills and knowledge" (2). TBLA conceives of assessment as a procedure that measures language learners' competence in conveying and comprehending meaning to achieve a particular end or outcome in authentic and communicative contexts (7, 8). In fact, a conspicuously vivid element in TBLA pertains to the explicit and direct assessment of the construct, i.e., performance (3).

Assessment Washback in EAP and TBLT

Assessment, as an ongoing process, is an integral component in educational systems of which language teaching and learning are no exception. In fact, assessment guarantees the accountability of instructional practices. Learning is not likely to come about unless appropriate and insightful approaches to assessment are adopted by teachers (7). Not unlike EAP assessment techniques, task-based tests entail learners to engage in certain tasks that are oriented towards the achievement of real-life outcomes (10).

Since EAP courses are essentially performance- or target language use oriented, the paradigm shift from DSAs towards performance assessment was ardently embraced by EAP specialists to get "a more valid construct of what it really means to know a language" (31, p. 188). Instances of performance tasks in an EAP context include note-taking during a lecture, writing an academic online forum, searching for and selecting relevant resources, giving an oral presentation, and writing a paper building upon multiple information sources (11).

Taking the advantages of TBLA for granted, such an assessment approach has robust repercussions on teaching and learning. In fact, education experts believe that testing exerts certain effects on teaching and learning. Such effect is termed as 'wash-back' (12), 'backwash' (13), or 'test impact' (14). Washback is defined as the direct and indirect impact of test on teaching and learning and drawing on what can be done in the classroom as a consequence of the test's effect, the washback can be positive and/ or negative (15). In practice, washback in itself appears to be neutral, however, as Ma (2021) argued, poorly constructed tests are more likely to result in negative washback (16).

A bulk of washback studies have been conducted since the publication of the remarkable work of Alderson and Wall (17) in virtually most levels of education across different disciplines, although there have been only a few washback studies (9, 18). The findings from the majority of these studies attribute negative washback effects to teachers' unfamiliarity and ignorance of curricular goals and to the gap between the curricular foci and those of the testing system. Gholami & Rajabi (2021) conducted a study to examine the effect of an evaluation cycle (Newly-developed Task Cycle) and Willis's task-based model on Iranian medical students' reading comprehension ability (19). To further explore the effect of TBLA on EAP students' reading ability, the following research questions were raised in the present study:

1. Does task-based language assessment significantly influence nursing students' reading comprehension ability?
2. Does non-task-based language assessment significantly influence nursing students' reading comprehension ability?
3. Is there any significant difference between the washback effects of task-based language assessment and non-task-based language assessment on nursing students' reading comprehension ability?

METHODS

3.1 Design of Study

The design of this study was pretest posttest non-equivalent groups. In this quantitative study, a quasi-experimental design was used to determine the washback impacts of task-based and non-task-based assessment on the Iranian nursing students' reading comprehension performance.

3.2 Participants

The participants of this research were 60 nursing students studying in the Nursing School of Kermanshah University of Medical Sciences. 41 students were female and 29 were male. They aged 19-24. They were selected from among the students of three intact classes (n= 78) on the basis of their performances on a quick Oxford Placement Test. The students who scored one standard deviation below and above the mean were chosen as the main participants of the study. These students (n=60) were non-randomly assigned to two experimental groups each comprising 30 learners. In the first group (i.e., Group1), task-based reading assessment was provided while in the second group (i.e., Group 2), non-task-based traditional reading assessment was administered. The groups signed a consent form as part of their tendency to take part in the study. They met two sessions in a week for about two months (16 sessions in total).

3.3 Instruments and Materials

The following instruments were employed in order to collect the required data for the present study.

a) The Quick Oxford Placement Test

The Quick Oxford Placement Test (QOPT) developed by Oxford University and University of Cambridge Local Examinations Syndicate (2001) is a standardized quick but still reliable and valid test to measure general language proficiency level. Due mainly to the ease of administration and other practical considerations, the QOPT was used as the yardstick to ascertain participating learners' homogeneity in terms of general English competency. The QOPT consists of 60 multiple-choice items on structure, reading comprehension and written expression. The mean and standard deviation of the participants' (N=60) scores on the QPT was M= 24.42 and SD= 3.86. According to CEFR (the European Common Framework of Reference), scores of 19-28 on this test belong to B1 (Pre-intermediate) level of language proficiency.

b) The Reading Comprehension Test

To gauge the participants' reading comprehension ability, the reading comprehension section of the Cambridge First Certificate in English (FCE, 2008) was used. This test comprises four parts: Part one, i. e., vocabulary, covers 8 items related to a modified cloze test containing eight gaps. For each gap, there was a multiple-choice item. Part Two, i. e., text structure, included seven questions based on a passage from which seven sentences were removed and placed in jumbled order right after the passage, together with an extra sentence which did not fit in any of the gaps. Part Three, i. e., morphology, encompasses eight items related to a text containing eight gaps. Each gap linked to one word. The missing word's root was provided after each gap in a parenthesis which had to be changed to provide the answer. Part Four, i. e., specific ideas, included seven items and exposes the testees to a long passage preceded by seven matching items. Testees had to locate the specific information which matches the items.

There were 30 items in the FCE in total. The FCE was available in two parallel forms. Accordingly, one form was used as the pretest and the other as posttest. At the end of the study, the scores from the two equivalent forms of the

FCE were correlated and the r equaled 0.824 which pointed to an acceptable reliability index of the test. The FCE is an internationally accredited test of English which is used widely by various business and educational institutions due to the fact it enjoys high psychometric properties. However, the reliability of the test was examined as follows. The test was piloted with a group of students with similar characteristics to the participants of the present study. The Cronbach alpha was calculated and the reliability of the test was 0.77. In addition to reliability, the validity of the test was also checked through administering it concurrently with the reading section of an actual IELTS test. The correlation coefficient index turned out to be 0.71.

c) Task-based and Non-task-based Assessment

Since the major concern in this study was to examine the washback effects of task-based and non-task-based assessment types on the participants' reading comprehension ability, two sets of tests were used as the main instructional foci of this study. The first set comprised task-based tests which were given to Group 1 participants every three sessions. These task-based reading tests were adopted from https://learnenglish.britishcouncil.org/skills/reading/b1-reading because of a number of reasons. First, these tasks have satisfactory psychometric indices (e.g., reliability and validity). Second, they are meticulously geared to the specific proficiency level of the target students. Third, they are free and easy to use. Examples of the task-based reading tests used with Group 1 are "A Conference Program, A Flyer for a Gym, and An email Request for Help". Each reading test consists of pre-reading activities and two reading tasks developed based on the passage. The second set of tests were teacher-made tests developed by the researchers to be administered to Group 2 students. These traditional tests consisted of matching, true/false, and open-ended items. The number of the items in each test varied as a function of the length of the reading passage. The participants in Group 2 took the non-task-based reading tests every three sessions.

Data Collection/ Analysis Procedure

The study took two months long. In the first session, the consent forms were signed by the students and the pre-test and proficiency test were administered followed by a brief introduction to the study and its objectives. In the same session, the participants were assigned to two experimental groups, i.e., Group 1 (n= 30) and Group 2 (n= 30). During the second session through to session 15, the participants received instruction on a specific academic English textbook, *English for the students of nursing* (48). Care was exercised to make sure that this resource would not teach to any of both task-based and non-task-based tests. Group 1 participants took a task-based reading test every three sessions and Group 2 students took a non-task-based reading test every three sessions.

There were five tests altogether in each group. In the last session, the participants in both groups took the reading post-test. Afterwards, the collected data were subjected to statistical analyses. Descriptive and inferential (Independent Samples *t*-test) statistics were run to analyze the data. Since in the first and second research questions, the aim was to

compare the same group's means on the pretest and posttest (within group differences), it was not appropriate to run ANOVA. Additionally, MANOVA could not have been used in that there was only one dependent variable, i. e., reading comprehension ability.

RESULTS

The results of statistical data analyses are presented below. Descriptive statistics of Group 1 and Group 2 participants' scores in pre- and post-tests indicated that the mean of Group 1 participants' scores on the reading comprehension pretest (M = 13.70) was less than that of Group 2 participants on the same test (M = 14.76). Conversely, the mean of Group 1 participants' scores on the reading comprehension posttest (M = 17.40) was more than that of Group 2 participants on the same test (M = 15.23).

To ascertain the normality of the obtained scores, One-Sample Kolmogorov-Smirnov test was run. According to the results, p values (p = 0.213, 0.233, 0.127, 0. 077) were higher than the critical level of significance (p = 0 .05) (*p > 0.05). Therefore, it was concluded that the data were normally distributed.

To answer the research questions, different statistical procedures and tests were run on the data. The first research question inquired if task-based language assessment significantly influences nursing students' reading comprehension ability. The pre- and post-test performances of Group 1 participants were compared using paired samples t-test.

The mean and standard deviation (SD) of Group 1 scores on reading comprehension pretest were 12.70 and 1.80 respectively. The mean and standard deviation (SD) of Group 1 scores on reading comprehension posttest were 17.40 and 1.16. Paired samples t-test was used to compare the means of Group 1 in reading comprehension pre- and post-tests. The results indicated that p value was less (t = -6.69, df = 29, two-tailed p = .00) than level of significance (p = 0.01) (*p < 0.01, two-tailed). Accordingly, there was a statistically significant difference between the pretest and posttest reading comprehension scores of Group 1 (task-based language assessment) which corroborated the positive impact of TBLA on nursing students' reading comprehension ability. That is, TBLA contributed to the improved reading comprehension performance of Group 1 participants from pretest to posttest.

The second research question inquired if non-task-based language assessment significantly influenced nursing students' reading comprehension ability. The pre- and post-test performances of Group 2 participants were compared using paired samples t-test. The mean and standard deviation (SD) of Group 2 scores on reading comprehension pretest were 14.76 and 2.14 respectively. The mean and standard deviation (SD) of Group 2 scores on reading comprehension posttest were 15.23 and 1.88. The results of paired samples test of Group 2 performance on reading comprehension pre- and post-tests disclosed that the p value was less (t = -3.50, df = 29, two-tailed p = 0.00) than level of significance (p = 0.01) (*p < 0.01, two-tailed). Therefore, the results indicated that there was a statistically significant difference between the

pretest and posttest reading comprehension scores of Group 2 (non-task-based language assessment) which pointed to the rather positive impact of non-task-based language assessment on nursing students' reading comprehension ability. That is, this type of assessment resulted in the improved reading comprehension performance of Group 2 participants from pretest to posttest.

The third research question read as: Is there any significant difference between the washback effects of task-based language assessment and non-task-based language assessment on nursing students' reading comprehension ability. To test the null hypothesis associated with this question and to examine the washback effects of task-based language assessment as the first independent variable and non-task-based language assessment as the second independent variable on nursing students' reading comprehension ability (i.e., dependent variable) in the pre- and post-test administrations, ANCOVA was used. The results of Levene's test ran to evaluate the equality of variances in reading comprehension post-test indicated that since the Sig value obtained (sig = .148) was more than the P value (.05), it could be concluded that the variance of the two groups were equal.

Based on the results of the analysis of homogeneity of the regression slope, since the obtained (p = .121) (F = 2.13) was more than α = .05, the null hypothesis was not maintained and the assumption of regression slope was observed. The covariance analysis, controlling pretest performances, indicated that there was a statistically significant difference between the washback effects of task-based and non-task-based language assessment on nursing students' reading comprehension ability (P < .05). As the F value (F = 28.64) at df = 1 and df = 59 was more than its critical value, thus the washback effects of task-based language assessment improved nursing students' reading comprehension performance as much as 33 percent compared to their pretest performance.

DISCUSSION

The present study set out to examine the washback effects of TBLA and NTBLA on nursing students' reading comprehension performance. The results of the study pointed to a statistically significant difference between the washback effects of the two assessment types. The results of the analysis of the data related to the first research question revealed that the washback effects of TBLA significantly improved the reading comprehension ability of the participants. This finding is in line with the one in Zoghi and Shahab's (2014) study where they reported the beneficial impact of TBLA on the participants' reading comprehension ability (20).

The findings are also in agreement with those from Abdollahi and Izadpanah (2021) in which the washback impact of TBLA was shown to significantly influence learners' vocabulary learning and grammatical ability (9).

The second research question investigated if the washback effect of NTBLA significantly affect the nursing students' reading comprehension ability of the participants. The results of the comparison of Group 2 performances on

reading comprehension pretest and posttest revealed that NTBLA significantly improved the students' reading ability. This finding lends support to the positive washback effect of traditional testing, i.e., NTBLA on the reading performance of the EAP students. This finding is in line with the findings from a number of previous washback studies (14) in which traditional NTBLA proved to be influential on casting certain levels of washback on students' language learning with particular reference to the learning of the four major skills. Compared to TBLA washback effects, it has to be stressed that NTBLA exerts a relatively lower level of washback an observation that sounds justifiable as the present researchers considered the exclusive features of TBLA including authenticity, integrity, motivation, and directness.

The third research question explored the differential washback effects of TBLA and NTBLA on reading comprehension of EAP students. The findings indicated that TBLA had a substantially more significant washback effect on nursing students' reading comprehension ability than NTBLA did. This finding, to reiterate, has to do with the peculiar characteristics of TBLA.

CONCLUSION

This study investigated the washback effects of TBLA and NTBLA on reading comprehension of Iranian nursing students. The results of the study revealed that both assessment approaches resulted in the improved reading comprehension performance of the participants with the former exerting a more influential impact than the latter. Furthermore, there was a significant difference between TBLA and NTBLA approaches with regard to their washback effect on the subsequent learning of reading comprehension of EAP learners.

A number of implications sound feasible in relation to the results of the study. First, EAP instructors and researchers should ponder on their assessment cultures and practices verify if and how their assessment practices help advance the students' language learning. Second, the chief goal behind the majority of assessment practices is to further educational development and learning, and owing to the effectiveness of TBLA in achieving such academic development and learning. This stands for reason to replace the traditional assessment techniques with TBLA as one alternative mode of assessment.

LIMITATION

This research is bound by a number of limitations and delimitations. The first limitation of the study is that the population involved in the investigation was confined to be nursing students at Nursing School of Kermanshah University of Medical Sciences. Second, given the number of variables and participants in the study, it was rather impracticable to triangulate the data through other measures. One of the delimitations of the study is that the participants were only adult students between the ages of 20 to 28 years old. The second delimitation was that it has been done in Iranian educational environment. Thus, the collected data were adequate only for describing perceptions of washback effect of task-based assessment.

REFERENCES

1. Mislevy R, Steinberg L, Almond R. Design and analysis in task-based language assessment. Language Testing, 2002; 19(4): 477-96.
2. Brindly G. 'Task-centered assessment in language learning: The promise and the challenge' in Bird N, Falvey P, Tsui A, Allison D, McNeill A. (eds.): Language and learning. Hong Kong: Hong Kong Institute of Language in Education, Hong Kong Education Department. 1994.
3. Long M.H, Norris J.M. Task-based language teaching and assessment. In Byram M. (Ed.), Encyclopedia of language teaching. 2000; 597-603. London: Routledge.
4. Ashrafi A. The effect of the task-based language teaching method on the comprehension ability of Iranian high-school students. International Journal of Multicultural and Multi religious Understanding, 2020;7(2): 154-60.
5. Chalak A. The effect of task-based Instruction on reading comprehension of Iranian EFL learners. Applied Research on English Language, 2015; 4(1): 19-30.
6. Dreyer C, Nel C. Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. System, 2003;31(3): 349-65.
7. Norris J.M, East M. Task-based language assessment. In Ahmadian M.J, H L.M, (Eds.). The Cambridge handbook of task-based language teaching. Cambridge University Press. 2021; 507-28.
8. Norris J.M. Task-based language assessment aligning designs with intended uses and consequences. JLTA Journal, 2018; 21: 3-20.
9. Abdollahi M, Izadpanah S. The washback impacts of task-based assessment on Iranian EFL learners' vocabulary learning and grammatical ability. International Journal of Research in English Education, 2021; 6(3): 21-39.
10. Khabbazi B, Anari F, Bagheri M.S, Sadighi F. An attitude-based survey of the impact of task-based assessment on the Iranian EFL learners' general English achievement. Research in English Language Pedagogy, 2020; 8(1): 101-22.
11. Spector-Cohen, E. Integrating performance assessment in the EAP classroom. The Internet TESL Journal, 2007; XIII(3): 133-37.
12. Wall D, Alderson J.C. Examining washback: The Sri Lankan impact study. Language Testing, 1993:10: 41-69.
13. Hughes A. Testing for language teachers (2nded.). Cambridge: Cambridge University Press. 2003.
14. Bachman L, Palmer A. Language testing in practice: Designing and developing useful language tests. Oxford: OUP. 1996.
15. Dong M, Fan J, Xu J. Differential washback effects of a high-stakes test on students' English learning process: evidence from a large-scale stratified survey in China. Asia Pacific Journal of Education, 2021; 3(1): 1-18.
16. Ma H. Washback effects of the IELTS test: views and experiences of Chinese students in the context of a Sino UK joint

program with English as the medium of instruction. [Doctoral dissertation, Queen's University Belfast]. 2021.

17. Alderson J.C, Wall D. Does washback exist? Applied Linguistics, 1993; 14: 115-29.

18. Al Amin M, Greenwood J. The examination system in Bangladesh and its impact: On curriculum, students, teachers and society. Language Testing in Asia, 2018; 8(1): 1-18.

19. Gholami M, Rajabi P. Effect of offline newly-developed task cycle (NDTC) on Iranian medical students' reading comprehension performance. International Journal of Foreign Language Teaching and Research, 2021; 9(37): 195-203.

20. Zoghi M, Shahab S. Task-based reading assessment vs. traditional reading test. Indian Journal of Fundamental and Applied Life Sciences, 2014; 4(S4): 1118-23.