

Negative Wash-back Effect in Online Assessment: Learners Inadequacy in Language Skills and Disengagement are the offspring

Omar Abu Al-majd (aasseell3@gmail.com)

Ph.D in Islamic Studies, Prince Sattam Bin Abdulaziz University, College of Education,
Department of Islamic Studies, Saudi Arabia

<https://orcid.org/0000-0001-9302-2003>

Abstract

The term wash-back effect describes how testing may affect several facets of education. The decisions that learners make and the choices that instructors make may both be influenced by testing. Wash-back may be seen as detrimental to more flexible techniques in language education, particularly in contexts where definitions of linguistic competence may be restricted; nevertheless, it may be seen as advantageous when strong teaching practices are the consequence. The current research intended to highlight the results of the negative washback effect in online instruction on learners' language achievement and engagement. To achieve this goal, 96 English as foreign language learners were grouped into an experimental group (EG) and a control group (CG). In both online instruction, language skills are taught at the intermediate level. The CG only intended to acquire language proficiency, whereas students in the EG aimed to pass a national exam. The pretests and posttests used the SInAPSi Academic Engagement Scale (SAES) and researcher-made tests to gauge learners' language proficiency. According to data screening, the language improvement of EG was less than CG. Moreover, the EG experienced more disengagement in online instruction and assessment. The results of this study have significant repercussions for the learning-oriented evaluation in online educational environments.

Keywords: Negative Wash-back Effect, Online Assessment, Language Skills, Disengagement, EFL Learners

Introduction

The rapid development and broad use of technological advances in communication and information have had far-reaching effects on many facets of human life, including the educational system. The innovations in technology have had a profound impact on our daily routines, the things we do, and the social standards to which we adhere. Similar to how technological progress has positively affected education in general, it has also impacted the teaching, learning, and testing of foreign languages (Shadiev et al., 2023; Song et al., 2023). Computer-assisted language learning (CALL) has come into being as a consequence of this transition, illustrating the potential of the World Wide Web for implementation in language instruction (Shadiev & Dang, 2022). With this shift came the rise of mobile-assisted language learning (MALL) as a popular means of communication, allowing EFL instructors and their students to connect with native speakers, create and begin sharing an infinite amount of information and subject matter, and interact in a virtual world without regard to the physical location or time of day (Hwang et al., 2022). Numerous studies on the effects of social media on EFL education have yielded the

conclusion that CALL and MALL provide EFL/ESL educators and students with access to meaningful and authentic language use, language learning input, unrestricted relationships, educational materials, and programs for students to utilize as well as assess proficiency in languages electronically.

The use of online assessments has grown increasingly significant in education not only because they are more relevant in virtual courses but also because they allow instructors to evaluate their pupils' progress in real-time. Both factors contribute to the fact that online assessments have been more widely used. In-class usage of digital assessment tools, such as those found online, helps to foster the growth of a more technologically advanced educational setting. Exams of this sort that are completed online provide educators a better opportunity to monitor their students' levels of interest and their progress.

Language competencies are becoming more vital in a globalized society, and knowing only one language is no longer beneficial. The goal of language learning should be to foster the growth of several skills, and participation in this process may help achieve this goal (Hwang et al., 2023). When exposed to a novel teaching style, students can engage more deeply in the learning process, emotionally invest in the material, and apply more effort to skill acquisition.

The term washback has been defined in various ways across the literature. Washback is defined quite simply by Shohamy et al. (1996) as the relationship between assessment and learning. In his seminal work, Gates (1995) characterizes it as the implication of assessment on education and instruction. More specifically, as Shohamy (1992) put it, washback is the use of outside language tests to influence and push foreign language learning...this condition is the result of the substantial power of extremal testing and the significant effect it has on the daily activities of test takers. According to Taylor (2005), washback dictates how, when, where, and what information test takers are instructed on. In this sense, Washback refers to the unintended consequences that tests may have on classroom instruction and student learning (Andrews et al., 2022).

The Saudi Arabian education system is heavily based on standardized testing. Education based on outcomes serves as a yardstick against which the achievement of university students in Saudi Arabia may be measured. Preparatory Year Program (PYP) courses in English are offered in Saudi universities. This curriculum is essential because it influences students' choices of majors in college, which in turn affects their professional development. Teachers in Saudi Arabia are expected to help their children do well on the PYP examinations, which are significant summative assessments. Some professors teach to the examinations instead of the students, a practice known as "harmful washback" because of the intense pressure to generate outstanding scores and the reality that PYP exams help students achieve their objectives. Therefore, it is not unusual for teachers to use all means required to ensure their pupils do well on tests.

Accordingly, Alqahtani (2021) found that the PYP significantly impacts the teaching processes, teaching staff attitudes and motivation, and assessment material across various school contexts. The exam's structure may be worked into classroom activities, more time and attention can be devoted to teaching students vocabulary and grammar, and innovative teaching strategies can boost students' grasp of course material. Teachers have taken to using quizzes as a means of helping students prepare for tests. These initiatives are designed to improve kids' prospects of doing well on standardized tests. However, there has been some negative washback on the curriculum due to the

exam's impact on teaching procedures. Some of these include putting more pressure on teachers and reducing their classroom flexibility (Al-Hinai & Al-Jardani, 2020).

Teachers are encouraged to develop further their methods of teaching English and to provide additional study resources to help their students do better on the test. The exam also motivates educators to adjust their practices to reflect students' needs better. The exam impacted the number of classes offered and the development of teaching materials. Some educators have narrowed their curriculum to focus only on exam preparation (Alqahtani, 2021). Like other high-stakes tests, the Learning Outcome-Based English Language Assessment (LOBELA) has far-reaching effects on tutoring and educational opportunities in Saudi Arabia. The LOBELA, like other rigorous examinations, has a far-reaching impact on tutoring and educational opportunities in Saudi Arabia.

It is worth highlighting that LOBELA is a crucial test in Saudi Arabia since the learners' future employment prospects depend on it. Teachers have the same responsibility as students to ensure they have the tools necessary to succeed on tests. Therefore, the criteria of this exam apply to instructors' pedagogical practices, content evaluations, and motives and attitudes. Therefore, many professors go out of their way to ensure their students are well-prepared for a LOBELA or similar high-stakes test (Alqahtani, 2021). As a result of participant feedback, testers often make curricular adjustments. One example of focusing on the essentials to improve students' grades is provided by Al Hinai and Al Jardani (2020). The negative impacts of testing may also cause teachers to cover topics that will appear on students' exams while ignoring the rest of the material covered in class (Allen, 2016).

EFL teachers are more likely to be affected by knowing which tests their pupils intend to take (Zhan & Andrews, 2014). Consequently, educators may change how and what they teach to better prepare students for the exam. Exam time is a good time to review material with pupils and ensure everything has been addressed. Exam preparation resources, such as sample questions and practice examinations, may be introduced by instructors to help students become used to the format of impending assessments. Students' reading habits and interests seem to be influenced by washback as well (Cheng et al., 2015).

Teaching the curriculum in positive washback is the same as teaching to the test. A negative washback effect arises when there is a discrepancy between what is intended to be taught and what is being tested. This might cause teachers to neglect classroom objectives instead of cramming for exams. In other words, if the test's content or style is predicated on a restrictive definition of language competence, it may have a chilling effect on the classroom environment and the students' learning capacity (Taylor, 2005). Overall, negative washout has more negative effects on education, especially language instruction, than positive ones. It limits the freedom of classroom instructors, students, curriculum designers, parents, and everyone else involved.

Academic engagement (AE) describes how much effort and time students put into completing classroom assignments and activities (Sharma & Bhaumik, 2013). According to the research of Fredericks et al. (2004), students' levels of academic engagement may be best explained as an outward reflection of their intrinsic drive to learn and succeed. It is widely agreed that academic engagement is a fluid term influenced by many factors inside and beyond the classroom (Shu, 2022). Findings suggest that students' levels of participation in L2 courses may be affected by their use of social media and other forms of technology. Students' personality and cognitive efforts to hone the material and abilities

linked with the course of study constitute the academic component of engagement (Nurjamin et al., 2023). Agentic engagement, the second facet of student participation, highlights the significance of students' efforts to enhance their own learning experiences. Students' social engagement is emphasized by extracurricular activities designed to foster interpersonal skills and problem-solving techniques (Ritonga et al., 2023). However, the idea of learner disengagement in assessment, which is the subject of the present study, was not correctly defined and analyzed in the academic work that came before it, and there is an obvious need for greater exploration into this issue.

Few empirical research efforts have attempted to determine how the negative washback effect affects language learning skills in L2, and to the best of our knowledge, no study has ever tried to investigate the impact of the negative washback effect on disengagement in online language learning and assessment. This research, therefore, intends to examine the influence of the negative wash-back effect in online assessment and their progress in language skills and AE in the context of EFL. The following research question was developed to help reach the goal of this study:

RQ: Does Public online Assessment (i.e., LOBELA) lead to negative wash-back, inadequacy in language skills, and disengagement in the Saudi Arabian EFL context?

Significance of this Research

To the best of the researchers' knowledge, this study is one of just a few that have focused on washback at the university level. Furthermore, the current research may substantially add to online testing and feedback, providing insight into the connection between testing and language education.

Methodology

Research Design

The present investigation is quantitative in scope, employing a quasi-experimental research style. The EG learned the four primary skills without considering LOBELA for one semester (16 sessions). Students in the CG (the class running concurrently) were allowed to be prepared for LOBELA. No LOBELA-based adjustments were made to the EG's instruction of the four focal skills. The teaching and evaluation of the two groups were online. The Oxford Quick Placement Test (OQPT) evaluated the students' English proficiency. With a possible range of 0.1 to 0.9 on the OQPT, a score of between 0.4 and 0.6 indicates moderate command of the English language. This research found a reliability of 0.89 for the OQPT.

Participants

Based on their scores on the OQPT, 96 students (55 females and 41 males) were selected from a pool of 173 first-year EFL students at the University of Jeddah, Saudi Arabia. These students had intermediate English language skills. They also took no additional English lessons before or throughout the study. Therefore, at the start of the research, all participants had an equivalent level of English ability.

They ranged in age from 19 to 23 and came from various socioeconomic backgrounds. The EG consisted of 47 students (19 female and 28 male), whereas the control group included 49 students (22 female and 27 male). The first semester of the university year was dedicated to required classes in all four sub-skills (i.e., speaking, listening, reading, and writing). Students provided their informed permission after being made aware of the nature of the study and its optional nature.

Materials

Pre-and Posttests

To delve into the participants' progress, a researcher-made test was developed and administered to the participants as pre and posttests. This test included two sections: one to assess the learners' academic development and the other to evaluate the AE. A LOBELA Mock exam was given to the students as a pretest to determine their proficiency in the test's four significant abilities. To assess the efficacy of treatment and the students' growth, a posttest consisting of the identical LOBELA Mock exam was administered at the conclusion of the trial. Moreover, to gauge the participants' engagement/disengagement, the Academic Engagement Scale (AES) was utilized. This instrument formulated by Freda et al. (2021) comprises six dimensions on a five-point Likert scale as follows: (1) the value of the university and a sense of belonging (six items), (2) the perception of the capability to persist in the university choice (four items), (3) the value of the university course (seven items), (4) engagement with university professors (four items), (5) engagement with university peers (five items), and (6) relationships between the university and relational networks (three items). According to the findings of this inquiry, the Cronbach alpha was 0.86, which suggested that the dependability was satisfactory.

Procedures and Data Analysis

Following the pretest, instructions were provided by one of the researchers, who also served as the language teacher for all of the classes taken by either the EG or CG participants. The students assigned to the CG received the standard curriculum, and no additional or supplemental material was used in their education beyond the primary textbook. Online instruction for the CG was intended to develop learners' language proficiency. For EG, the main objective was to prepare them to achieve the required qualifications to pass LOBELA. The posttest was administered after this project and after the teaching had been finished. The test's purpose was to evaluate the students' accomplishments and AE in both the CG and the EG and to determine how successful the program had been.

Subsequently, Independent and Paired Samples T-tests were performed to look at the negative wash-back effect in online teaching and assessment and its implications for learners' deficiency in language skills and academic disengagement.

Findings

Firstly, the normality of the distribution was checked by a robust statistical test (i.e., Kolmogorov-Smirnov), the results of which are shown in Table 1.

Table 1

Kolmogorov-Smirnov Results for the pretests and posttests

	Kolmogorov-Smirnov		
	Statistic	df	Sig.
EG LA Pretest	.092	47	.20

EG LA Posttest	.141	47	.09
EG AE Pretest	.138	47	.08
EG AE Posttest	.166	47	.11
CG LA Pretest	.113	47	.174
CG LA Posttest	.102	47	.20
CG AE Pretest	.180	47	.07
CG AE Posttest	.146	47	.21

EG: Experimental Group; CG: Control Group; LA: Language Achievement; AE: Academic Achievement

In Table 1, the *Sig.* value under the Kolmogorov-Smirnov part of the table showed a value lower than .05 ($\text{Sig.} > 0.05$), which indicates that the distribution of scores for the pretests and posttests of LA and AE words was normal. Therefore, using parametric statistics such as independent and paired samples *t*-tests is reasonable.

Secondly, to ascertain the homogeneity of the two groups in terms of their LA and AE before the treatment, their pretest scores were compared via an independent-samples *t*-test:

Table 2
Descriptive Statistics for the LA and LE Pretests

	Groups	N	Mean	Std. Deviation	Std. Error Mean
LA Pre	EG	47	33.5532	9.54579	1.39240
	CG	49	31.0204	11.84886	1.69269
AE Pre	EG	47	21.7660	3.81195	.55603
	CG	49	20.1837	5.99887	.85698

Table 2 shows that the EG learners' mean score on the LA and AE equaled 33.5532 and 21.7660, respectively, and the CG learners' mean score was 31.0204 and 20.1837. To determine whether the difference between these two mean scores and, thus, the two groups on the pretests was statistically significant, the researcher had to examine the *p*-value under the *Sig.* (2-tailed) column in the *t*-test table. In this table, a *p*-value less than .05 would indicate a statistically significant difference between the two groups, while a *p*-value larger than .05 suggests a difference that failed to reach statistical significance.

Table 3
Independent Sample T Test (LA and AE Pretests)

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference
LA Pre	Equal variances assumed	3.387	.069	1.150	94	.253	2.53278	2.20164
	Equal variances			1.156	91.311	.251	2.53278	2.19180

AE Pre	not assumed Equal variances assumed	.586	.446	1.535	94	.128	1.58228	1.03074
	Equal variances not assumed			1.549	81.794	.125	1.58228	1.02156

Based on the information in Table 3, there was no statistically significant difference in the LA and AE pretest scores for CG ($p > 0.05$). Hence, it could be inferred that the learners in the two groups were at the same level of LA and AE.

The study's research question was to determine whether Public online Assessment (i.e., LOBELA) leads to negative wash-back, inadequacy in language skills, and disengagement in the Saudi Arabian EFL context. To find an answer to this research question, the pretest and posttest scores of the learners in the EG and CG were compared utilizing a paired-sample *t*-test:

Table 4

Descriptive Statistics (EG and CG pretests and Posttests)

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	EG LA Post	29.9362	47	15.06999	2.19818
	EG LA Pre	33.5532	47	9.54579	1.39240
Pair 2	EG AE Post	19.2979	47	3.97774	.58021
	EG AE Pre	21.7660	47	3.81195	.55603
Pair 3	CG LA Post	42.7755	49	10.97775	1.56825
	CG LA Pre	31.0204	49	11.84886	1.69269
Pair 4	CG AE Post	25.8163	49	3.14691	.44956
	CG AE Pre	20.1837	49	5.99887	.85698

It could be observed in Table 4 that the difference between the pretest and posttest mean scores of both groups on LA and AE was quite substantial. To find out whether this difference was statistically significant or not, the following *t*-test table had to be checked:

Table 5

*Results of the Paired-Samples *t* Test (EG and CG pretests and Posttests)*

		Mean	Std. Deviation	Std. Error Mean	t	df	Sig.
Pair 1	EG LA Post – EG LA Pre	-3.61702	11.91009	1.73726	-2.082	46	.033
Pair 2	EG AE Post – EG AE Pre	-2.46809	2.80412	.40902	-6.034	46	.000
Pair 3	CG LA Post – CG LA Pre	11.75510	11.90156	1.70022	6.914	48	.000
Pair 4	CG AE Post – CG AE Pre	5.63265	5.03196	.71885	7.836	48	.000

Table 5 revealed that there was a statistically significant difference between the LA pretest and posttest scores of the EG learners since the $p < .05$). Moreover, the difference between the AE pretest and posttest scores of the EG learners was significant, divulging that the EG LA and AE decreased after the treatment. This indicates that online

assessment (i.e., LOBELA) leads to negative wash-back on LA and LE in Saudi Arabian EFL.

Discussion

This research aimed to investigate the potential effects of negative washback on subskills (i.e., speaking, listening, reading, and writing) among EFL university learners in Saudi Arabia. According to the findings, placing more of a focus on high-stakes exams and the consequences that they may have in future life of the students may hinder the language competency of the learners. Teaching to the test may also affect EFL learners' engagement in virtual instruction and assessment. Compared to the control group, which showed some progress in language learning, the EG demonstrated some improvement, but it was not statistically significant. Due to the negative washback effect, EG experienced disengagement in online instruction and assessment.

The findings of this research indicated that negative washback is unavoidable when the stakes are high. Instruction that prepares students for high-stakes exams is the principal objective of limiting how teachers can instruct and diminishing the affordance available to instructors. EFL teachers restrict their selections of teaching topics, methods, and materials and (b) motivate students to use a memorization approach instead of critical thinking. Another explanation for these outcomes is that standardized tests can place educators in an impossible situation in which they are forced to give up their autonomy and freedom of choice regarding the curriculum and instruct what they should be teaching rather than what they prefer to teach. This finding is supported by what Razavipour et al. (2018) and Cheng et al. (2015) concluded.

According to the findings of Jamalifar et al. (2021), altering the language learning and teaching process necessitates changing both the content and structure of high-stakes exams. They went on to say that the substance of the examination should be geared toward real-world scenarios that need communication to lessen the test's impact on students. In the same vein, Zhan and Andrews (2014) concluded that the large-scale TOEFL iBT exam had an influence on the content as well as the technique of instruction but that these effects were mediated by the usage of preparation resources for the test. In the words of Puspitasari and Influx (2023), a thorough knowledge of washback in the teaching and learning process is crucial to making appropriate interventions. Once the reasons have been correctly identified, the author argues that relevant improvements, such as teacher training or exam redesign, may be implemented.

The premises of social-cognitive theory (Bandura, 2012) provide credence to this result by emphasizing the significance of students' active participation in self-regulation and reflection and developing more positive viewpoints. It is implied that EFL students will feel more involved and less worried during online evaluation if they have practiced effective ways of conducting their exams and shown their genuine ability. This may be achieved by learning how to optimally handle the period immediately before, during, and after an exam. The goal may be attained by teaching the students specific valuable methods. They may learn to calm their nerves, keep track of time, and do well on the test by just taking practice exams. In line with the results of the current study, Ritonga et al. (2023) discovered that engagement in online assessment may be accomplished by fostering test-taking skills, resilience, and autonomy in students.

This finding is consistent with the motivation/demotivation theories (Ryan & Deci, 2000). Practicing higher-order thinking skills may be helpful for EFL students since it allows them to reflect on their feelings and develop creative strategies for dealing with

the anxiety and challenges that come with forthcoming online exams. According to the self-determination hypothesis (Martin & Marsh, 2009), motivation, contentment, and classroom enjoyment all benefit from an increased understanding of one's strengths and weaknesses. It seems sensible to assume that students' language skills will increase if they are given more freedom to study languages on their terms via online courses. Ismail and Heydarnejad (2023) came to similar conclusions on the connections between self-efficacy, linguistic ability, and optimal performance.

This research suggests that fostering a trustworthy and encouraging environment in language learning might help students feel more comfortable participating in class activities. It can also be inferred that university students more committed to their learning goal may be more likely to stick with it over the long haul. They can more accurately evaluate themselves if they are consistent in their activities and plans. Previous studies (Zhao & Liao, 2021) have demonstrated that learning a language is a cognitive activity that calls for the participants to be actively interested and engaged in the process. Learners in higher education may find these difficulties more perplexing because of the unique nature of language acquisition. EFL university students with a healthy dose of positive self-concept and emotional health (Namaziandost et al., 2023) are likelier to use effective cognitive, metacognitive, and problem-solving procedures. Stakeholders' interest in standardized testing is on the rise because of its significant influence on classroom instruction. Taylor (2005) pointed out that students, educators, and other interested parties will continue to take an active interest in testing as long as it is employed on a widespread, global scale. Washback requires forethought, attention, research, and information dissemination, as Djuri (2015) stated.

Conclusion

The purpose of this research was to contribute to the barely body of information about the washback effect in online instruction and its likely impact on language abilities and AE among EFL students. The findings of this study, when taken as a whole, lead the researchers to the conclusion that it is essential to take into account the emotional experiences of language students, in addition to tracking their development in cognition and metacognition, to effectively teach and learn a foreign language to ensure their engagement in learning and assessment.

There is always a connection between instruction, assessment, and evaluation. These connections may be beneficial or harmful. The introduction of new testing procedures impacts teachers and children alike. Testing has a significant influence that should not be overlooked. It has the potential to either improve teaching and learning (positive washback) or derail current initiatives in the classroom. The test has a continuous, multi-layered impact. Student attrition and engagement may be affected by many things, but the environment, the assessment, and the instructor all play significant roles. Workshops and training should be provided for educators to improve their evaluation techniques. If they test themselves and monitor their progress, students are more likely to have a favorable outlook on the effect of examinations on their learning. This will help them succeed on the exam by giving them feedback on how they are doing.

It is strongly recommended that those involved in the evaluation process reflect on the test's results and make the appropriate local and global adjustments. Students should also choose the most effective study methods for themselves; teachers should

model and discuss test-taking tactics before and throughout the administration of achievement exams. Writing comments on students' test papers as an exam-taking strategy after administering mid-semester achievement assessments is one way of offering students helpful insight into their language learning progress. More importantly, positive washback lies not only with instructors but also with their students.

A positive washback is generated when a clear link is made between educational objectives and evaluation. This connection should be made clear to both students and instructors. They speculated that the continuous training on assessment techniques offered to teachers was likely responsible for the positive washback in teaching and learning. It is highly recommended that students be given assignments that enable them to develop practical self-aid constructions in addition to academic topics. Both personal tracking and self-awareness are skills that should be exercised from the first stages of learning a new language. Consequently, students will have a higher chance of obtaining academic performance, assessment will be more centered on their requirements, and society will benefit as a whole.

Some of the limitations of the current study may warrant further exploration in future research. Due to the quasi-experimental nature of the subject selection, these results have limited applicability. Longitudinal studies of the impacts of washback effects and AE among second-language students are required. Furthermore, demographic factors were not analyzed. Researchers are strongly encouraged to include demographic information on teachers in future studies. This research used a quantitative approach, and it is feasible that a mixed-methods investigation will provide better results.

Funding

"The author extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the Project number (PSAU/2023/02/25639)"

References

- Allen, D. (2016). Investigating washback to the learner from the IELTS test in the Japanese tertiary context. *Language Testing in Asia*, 6(1), 1-20. <https://10.1186/s40468-016-0030-z>
- Alqahtani, F. (2021). The Impact of Language Testing Washback in Promoting Teaching and Learning Processes: A Theoretical Review. *English Language Teaching*, 14(2), 21-26. <https://doi.org/10.5539/elt.v14n7p21>
- Al-Hinai, M., & Al-Jardani, K. (2020). Washback in language testing: An exploration with a focus on a specific EFL context in Oman. *International Journal of English Linguistics*, 11(1), 68-75. <https://doi.org/10.5539/ijel.v11n1p68>
- Amerstorfer, C. M., & Freiin von Münster-Kistner, C. (2021). Student perceptions of academic engagement and student teacher relationships in problem-based learning. *Frontiers in Psychology*, 12, 713057. <https://doi.org/10.3389/fpsyg.2021.713057>.
- Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, 4, 87. <https://doi.org/10.3389/educ.2019.00087>.
- Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback - A case-study. *System*, 30(2), 207-223. [https://10.1016/S0346-251X\(02\)00005-2](https://10.1016/S0346-251X(02)00005-2)

- Bandura, A. (2012). Social cognitive theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 349–373). Sage Publications Ltd. <https://doi.org/10.4135/9781446249215.n18>
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching*, 48(4), 436-470. <https://10.1017/S0261444815000233>
- Freda, M. F., Rafaele, D. L. P., Esposito, G., et al. (2021). A new measure for the assessment of the university engagement: The SInAPSi academic engagement scale (SAES). *Current Psychology*. <https://doi.org/10.1007/s12144-021-02189-2>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. <https://doi.org/10.3102/00346543074001059>
- Gates, S. (1995). Exploiting washback from standardized tests In J. D. Brown & S' O' Yamashita (Eds.), *Language testing in Japan* (pp' 101-106). Tokyo: Japanese Association for Language Teaching.
- Hwang, W.-Y., Guo, B.-C., Hoang, A., Chang, C.-C., & Wu, N.-T. (2022). Facilitating authentic contextual EFL speaking and conversation with smart mechanisms and investigating its influence on learning achievements. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2022.2095406>.
- Hwang, W.-Y., Nurtantyana, R., Purba, S. W. D., Hariyanti, U., Indrihapsari, Y., & Surjono, H. D. (2023). AI and recognition technologies to facilitate English as foreign language writing for supporting personalization and contextualization in authentic contexts. *Journal of Educational Computing Research*. <https://doi.org/10.1177/07356331221137253>
- Jamalifar, G., Salehi, H., Tabatabaei, O., & Jafarigoha, M. (2021). Washback Effect of EPT on Iranian PhD Candidates' Attitudes Toward Learning Materials. *International Journal of Foreign Language Teaching and Research*, 9 (39), 11-26.
- Namaziandost, E., Heydarnejad, T., & Azizi, Z. (2023). To be a language learner or not to be? The interplay among academic resilience, critical thinking, academic emotion regulation, academic self-esteem, and academic demotivation. *Current Psychology*. <https://doi.org/10.1007/s12144-023-04676-0>
- Nurjain, A., Salazar-Espinoza, D. E., Saenko, N., & Bina, E. (2023). Learner-oriented assessment matters: testing the effects of academic buoyancy, reflective thinking, and learner enjoyment in self-assessment and test-taking anxiety management of the EFL learners. *Lang Test Asia* 13, 30. <https://doi.org/10.1186/s40468-023-00247-z>
- Puspitasari, M., & Influx, M. A. P. (2023). The Mapping of Mediating Negative Washback of The National Examination. *Eltin Journal: Journal of English Language Teaching in Indonesia*, 11(1). <https://doi.org/10.22460/eltin.v11i1.p87-98>
- Razavipour, K., Moosavinia, S. R., & Atayi, S. (2018). Construct ambiguity and test difficulty generate negative washback: the case of admission test of English literature to graduate programs. *International Journal of Instruction*, 11(4), 717-732.
- Ritonga, M., Shaban, A. A., Al-Rashidi, A. H., & Chilani, N. (2023). Engagement in online language assessment: are test-taking skills, self-assessment, resilience, and

- autonomy critical? *Lang Test Asia*, 13, 25. <https://doi.org/10.1186/s40468-023-00236-2>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Shadiev, R., & Dang, C. (2022). A systematic review study on integrating technology-assisted intercultural learning in various learning context. *Education and Information Technologies*, 27(5), 6753–6785. <https://doi.org/10.1007/s10639-021-10877-6>
- Shadiev, R., Wen, Y., Uosaki, N. & Song, Y. (2023) Future language learning with emerging technologies. *Journal of Computers in Education*. 10, 463–467 (2023). <https://doi.org/10.1007/s40692-023-00285-9>
- Sharma, B. R., & Bhaumik, P. K. (2013). Student engagement and its predictors: an exploratory study in an Indian Business School. *Global Business Review*, 14(1), 25–42. <https://doi.org/10.1177/0972150912466364>
- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *Modern Language Journal*, 76, 513–521. <https://doi.org/10.2307/330053>
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13, 298–317. <https://doi.org/10.1177/026553229601300305>
- Shu, K. (2022). Teachers' commitment and self-efficacy as predictors of work engagement and well-being. *Frontiers in Psychology*, 13, 850204. <https://doi.org/10.3389/fpsyg.2022.850204>
- Song, Y., Wen, Y., Yang, Y., & Cao, J. (2023). Developing a 'Virtual Go mode' on a mobile app to enhance primary students' vocabulary learning engagement: An exploratory study. *Innovation in Language Learning and Teaching*, 17(2), 354–363. <https://doi.org/10.1080/17501229.2022.2047693>
- Taylor, L. (2005). Washback and impact. *ELT Journal*, 59(2), 154–155. <https://doi.org/10.1093/eltj/ccj030>
- Thaidan, R. (2015). Washback in language testing. *Education Journal*, 4(1), 5–8. <https://doi.org/10.11648/j.edu.20150401.12>
- Zhan, Y., & Andrews, S. (2014). Washback effects from a high-stakes examination on out-of-class English learning: Insights from possible self-theories. *Assessment in Education: Principles, Policy & Practice*, 21(1), 71–89. <https://doi.org/10.1080/0969594X.2012.757546>
- Zheng, X., Ismail, S.M., & Heydarnejad, T. (2023). Social media and psychology of language learning: The role of telegram-based instruction on academic buoyancy, academic emotion regulation, foreign language anxiety, and English achievement. *Heliyon* 9(5): e15830. <https://doi.org/10.1016/j.heliyon.2023.e15830>