
Revisiting Washback Inquires in Indonesia

MAYA PUSPITASARI*

Abstract

This research revisits the washback concept in Indonesia's education system, specifically focusing on the impact of national standardized tests on teaching and learning practices. In Indonesia, high-stakes testing, such as the national examination, has been used to assess educational quality, often influencing curriculum design, teaching methods, and student motivation. The research highlights how such examinations can narrow the curriculum, encourage teaching to the test, and place undue pressure on stakeholders, including teachers and students. Drawing from research conducted in Indonesia and globally, the research also discusses the need for further investigations to develop assessment models that promote positive washback effects. This shift opens opportunities for rethinking assessments that support holistic learning and skill development. The research concludes by emphasizing the importance of well-designed assessments in fostering positive educational outcomes and improving the quality of education in Indonesia.

Keywords

High-stakes testing, national examination, students, teachers

Article History

Received 06 October 2024

Accepted 19 April 2025

How to Cite

Puspitasari, M. (2025).
Revisiting washback inquires
in Indonesia. *Indonesian
Research Journal in Education
| IRJE |*, 9(1), 398-409.
<https://doi.org/10.22437/irje.v9i01.37467>

*Universitas Terbuka, Indonesia, Corresponding author: maya_p@ecampus.ut.ac.id

Introduction

The education system reinforces the quality of a country in place. Failure of schools, indicated by low scores in standardised high-stakes testing, may contribute to the country's slow progress toward improvement. Questions must be raised about this country's ability to sustain human development in the short and long term. A standardised test, such as the national examination, can be interpreted as an attempt to assess the quality of education in Indonesia. Even though education policy is constantly changing, policymakers may need to identify which aspects of education are the most important for a successful education system. According to the 2024 United Nations Development Programme (UNDP) report, Indonesia ranked 112th, with lower human development than its neighbours Singapore, Brunei Darussalam, Malaysia, Thailand, and Vietnam, as in the following table.

Table 1. *The Asian countries rank by the Human Development Index (HDI)*

Country	Rank	
	2021	2022
Singapore	10	9
Brunei Darussalam	53	55
Malaysia	68	63
Thailand	69	66
Vietnam	108	107
Indonesia	113	112

The government's expectation to improve education quality may only create a burdensome teaching and learning environment. The goal of effective learning is likely to narrow the curriculum to achieve high exam scores only. Berliner (2011) described the situation well. He stated that if student achievement tests carry significant consequences for teachers and administrators, the pressure can lead to unethical behaviour, such as cheating or disregarding test standardization procedures. Additionally, teachers often focus heavily on test preparation, undermining the validity of high-stakes testing. Educators may even manipulate student placement by removing underperforming students from their classes, encouraging them to drop out, or holding them back a year to improve their chances of performing better on the tests.

National test results frequently impact the success or failure of schools, teachers, and students. National exams, such as standardised assessments, can substantially impact on how schools, teachers, and students are assessed. These tests are frequently used to assess student performance and progress, and the results can be used to determine school funding, teacher evaluations, and even future educational and career opportunities for students. While these exams can help assess performance, it is important to consider multiple factors when assessing the success or failure of schools, teachers, and students because test scores do not always reflect the individuals' and institutions' full range of abilities and achievements. It is in line with Jeffrey & Geoff (2014), who emphasised that school success is defined simplistically in performative terms of specific proportions of students in a particular school achieving grades, which places demands on teachers' work and students. High scores can also lead to increased

enrollment and positive school publicity, whereas low scores can lead to decreased enrollment and negative publicity (Puspitasari, 2020). Scores can also assess the efficacy of a school's curriculum, teaching staff, and overall educational programme. It can also be used to identify areas where the school can improve and changes that can be implemented to support student learning and success better. Scores can also assess a school's performance and allocate funding and resources. This is where the washback effect might come into play.

The research outlines various research conducted, highlighting their relevance and the potential for replication within the Indonesian context. It then presents several studies from Indonesia aimed at generating interest in conducting similar research with different participants across diverse regions. Additionally, the research discusses discrepancies between the current and previous research conducted in Indonesia, emphasizing gaps in knowledge.

Methodology

This research was qualitative research with a descriptive approach. This research revisited the washback concept in Indonesia's education system, specifically focusing on the impact of national standardized tests on teaching and learning practices. This research also investigated the development of assessment models that promote positive washback effects. The research highlighted how such examinations can narrow the curriculum, encourage teaching to the test, and place undue pressure on stakeholders, including teachers and students. The research outlined various research conducted and highlighted their relevance and the potential for replication within the Indonesian context. Furthermore, presented several studies from Indonesia were presented to generate interest in conducting similar research with different participants across diverse regions. Additionally, the research discussed discrepancies between the current research and previous research conducted in Indonesia, emphasizing gaps in knowledge.

Findings and Discussion

Why should washback be an issue?

Initially introduced by Buck in 1988, scholars have extensively explored the concept to examine how language examinations shape attitudes, pedagogical practices, curricula, and student motivation. The washback effect has also been employed to investigate the broader implications of testing, extending beyond teachers and students to include parents, school principals, textbook publishers, test designers, and policymakers who indirectly influence the educational process.

Spratt (2005) defined washback as a test's impact on the teaching and learning done in preparation. This definition emphasizes the preparatory activities influenced by the nature of the test, highlighting how tests can shape both instructional practices and student behaviors. A test can have a washback on the teaching and learning process, both positively and negatively. The washback concept is significant because it emphasises the importance of designing tests in a way that aids rather than hinders language learning. Well-designed tests have been shown to have a positive washback effect by motivating students to learn and providing teachers with valuable feedback on student performance (Pitoyo & Asib, 2020). Students' motivation to study increases significantly when they know their academic future

depends on their performance. As [Mons \(2009\)](#) noted, testing influences student behaviour by increasing motivation to study." This suggests that high-stakes assessments can be a powerful external motivator, driving students to invest more effort and time in their studies. The pressure to succeed often leads students to adopt more focused and disciplined study habits, knowing that their test results may affect important academic decisions, such as graduation, advancement to the next level, or future opportunities. Consequently, testing plays a critical role in shaping students' learning behaviours and attitudes toward their education. On the other hand, poorly designed tests can have a negative washback effect by discouraging students and teachers and causing a narrow focus on test preparation rather than language learning.

Students are often highly motivated to avoid repeating a grade or taking remedial exams, which drives them to work harder to achieve the minimum scores required for high school graduation. The pressure is further amplified as exam results are frequently published in local and national newspapers and broadcast on various electronic media platforms, such as television and the Internet. This creates a heightened level of stress among not only students and teachers, school principals and parents. [Barksdale-Ladd and Thomas \(2000\)](#) argued that the publication of test score comparisons at various levels like school, district, and state in media, combined with intense pressure from educational authorities, including administrators, school boards, and legislators, places a heavy emphasis on test performance. As a result, teachers tend to focus their instructional strategies primarily on improving test scores rather than on broader educational goals. This external pressure makes testing a central concern in their teaching approach, shaping how and what they teach to meet the high-stakes demands of standardized assessments.

[Fernández and Martínez \(2022\)](#) emphasized that although standardized test scores indicate teacher effectiveness, they represent only one part of a much larger and more complex picture. Relying solely on test scores to evaluate teachers' or student performance overlooks other critical factors that significantly shape educational outcomes. For example, parental support can profoundly influence students' motivation and ability to succeed, as engaged parents often provide additional resources, encouragement, and structure that enhance learning. The quality of school facilities, such as access to modern technology, well-maintained classrooms, and adequate learning materials, can either support or hinder the educational process. Schools with better resources provide students with a more conducive learning environment, which can positively impact their performance.

Moreover, students' learning experiences inside and outside the classroom also affect their achievement on national exams. Engaging, interactive lessons and extracurricular activities can deepen students' understanding of key concepts and improve critical thinking skills, essential for academic success. Outside the classroom, experiences such as tutoring, community involvement, and exposure to diverse perspectives can further enrich students' educational journeys. Therefore, while standardized test scores can provide valuable data, they must be considered alongside these broader influences to understand and assess student performance and teacher effectiveness fully.

Washback effect on teachers and students

According to Madaus (1988), when test results are the sole or even partial arbiter of future educational or life choices, society tends to treat test results as the primary goal of schooling rather than as a useful but fallible indicator of achievement. There are likely reasons why teachers prefer to teach based on the accountability of stakeholders rather than helping students feel good about school and learning in the classroom. Teachers are obligated by society to assist students in facing the liability of the exam. Because the examination is a critical component of continuing their studies at a higher level, the students will most likely press the teachers to improve their exam skills. Goodson (2003) referred to teachers as key individuals who play an important role in ensuring that education is effectively delivered across various schooling systems. Teachers may have to tailor their professionalism to meet the needs of the stakeholders. A high-stakes test may prompt teachers to create exam preparation classes so students can become accustomed to high-stakes situations like those in exams.

Numerous studies worldwide have explored the washback effect of high-stakes tests on different stakeholders, including teachers and students, revealing how tests shape what and how teaching and learning proceed in the classroom. The research on the washback effect of the Cambridge English First Certificate Exam (FCE) in Spanish secondary schools reveals that teachers often adapt their instructional methods to align with the exam's content and format, leading to a more exam-focused curriculum (Tsagari, 2011). This adaptation can sometimes come at the expense of broader educational goals. The pressure of preparing for the FCE affects students' learning behaviors, with some students becoming more motivated and focused while others experience increased stress and anxiety. Additionally, the exam influences the design of the English language curriculum, often narrowing its scope to what is tested, which may limit the development of comprehensive language skills. Teachers may also feel constrained by the need to teach to the test, reducing their ability to implement diverse and innovative teaching practices. The research suggests that while standardized exams like the FCE provide valuable data on student performance, they should be used alongside other assessment methods to holistically evaluate teaching effectiveness and student learning.

A decade later, Dong et al. (2021) conducted quantitative research in China to examine the differential washback effects of the National Matriculation English Test (NMET) on the learning process. The research involved 3,105 senior high school students of various gender, grades, and English proficiency levels across southwestern China. The findings show that male students tend to be more test-driven in their motivation, while female students experience higher levels of anxiety and spend more time on learning practices. The research also reveals that the NMET influences students in higher grades regarding learning motivation and anxiety, although time spent on learning does not significantly change. Furthermore, students with higher English proficiency face fewer negative washback effects on motivation and anxiety yet demonstrate more consistent learning practices.

In Tehran Estaji and Alikhani (2020) delivered a questionnaire to 50 teachers and 50 students and conducted structured interviews with 10 teachers and 10 students from five private cramming schools. The results revealed that the teachers concentrated exclusively on preparing students for the exam, tailoring their lessons to meet the specific demands of the

First Certificate in English (FCE). This intense focus on exam preparation impacted the teachers, as their professional worth was tied to their students' exam results. Teachers reported that the FCE exam strongly influenced their teaching methods, materials, and exam preparation strategies, such as using past papers and supplementary materials. Meanwhile, students had positive perceptions, associating the FCE exam with their professional and educational advancement.

How the washback effect has been investigated in Indonesia

Sulistyo (2009) administered a yes-no questionnaire to 24 English teachers from Islamic lower secondary schools in seven Indonesian provinces. The questionnaire asked teachers about their thoughts on the relationship between the implementation of national exams and students' English proficiency. During their teaching practises, teachers tended to guide students in achieving high exam scores, according to the results. Teachers may successfully motivate students to work harder to prepare for the national exam. However, it was also revealed that the teachers saw the situation as an opportunity to hold an extra learning programme outside of school hours to earn more money. Sulistyo realised that more research was needed to thoroughly examine the impact of national examinations on the larger population. A single method he used also implied the need for additional research to determine how strong the influence of national examinations is in teaching practice. Sulistyo's findings might be explored further by conducting semi-structured interviews. Lichtman (2012) claimed that "participants can share what they know and have learned and can add a dimension to our understanding of the situation that questionnaire data or a highly structured interview does not reveal."

Sukyadi and Mardiani (2011) investigated the washback effect of the English national examination at a secondary school in Bandung using four methods: classroom observation, questionnaire, interview, and document analysis. The three secondary schools participating in the research were divided into low-, medium-, and high-level achievers based on their exam scores over the previous four years. Seven teachers who completed a questionnaire demonstrated that the national examination harmed the teaching methods used in the classroom. Sukyadi and Mardiani (2011) discovered that teachers tend to teach to the test by providing past test papers. The classroom activity focused on going over previous years' English national exams. Because the national examination has less of an impact on the learning process of first and second-year students, including them in the research will increase the number of participants.

A questionnaire was given by Aprianto (2013) to 19 English teachers at upper secondary schools to address the washback issue in the teaching and learning process. All respondents stated that the only thing they did in the classroom was prepare students to be able to answer the test and motivate them, to be honest when doing the test. There is a tendency for the teachers being investigated to disagree that the national examination was the only instrument used to determine students' graduation. According to the respondents, the students were burdensome because they needed specific scores on the national exam to graduate and continue their studies at a higher level.

The researchers set out different approaches to investigate the washback effect of the national examination in which the conclusion meets an agreement, drawing on a wide range of sources. The researchers agreed there should be a way to evaluate students' performance without causing a negative washback effect. However, none of the studies mentioned above clearly suggested a further investigation to offer a specific model to create a positive washback effect in the teaching and learning process that might meet the teachers' expectations. Sulisty (2009) reckons a strong desire to implement classroom-based assessments, which focus on evaluating students' learning directly within the classroom through various methods such as quizzes, projects, and observations. Instead of standardized testing, this type of assessment places greater responsibility on teachers. They are expected not only to teach but also to design, administer, and evaluate these assessments. Aprianto (2013) demonstrated that teachers desired an improved testing system for assessing students' competencies by referencing the idea of assessment that actively supports the learning process. His research highlighted teachers' preference for a more effective and supportive assessment system. They are seeking a method of evaluation that measures students' competencies and enhances their learning experience. By referencing the concept of assessment that supports learning, Aprianto suggested that teachers are advocating for a system where assessments are integrated into the learning process, providing constructive feedback and promoting student growth rather than just serving as a final judgment of their abilities.

In recent years, the educational landscape in Indonesia has changed dramatically, particularly with the abolition of the national examination 2016. This watershed moment has reignited attention to the washback concept, which refers to the impact of testing on education. With the withdrawal of the UN, educators and policymakers are now investigating how this move affects classroom practices, teacher attitudes, and student learning results. The focus has switched to high-stakes assessments such as the university entrance exam (SBMPTN), which continues to impact the educational process significantly. This developing scenario provides a unique chance to investigate the positive and negative washback effects in the Indonesian setting, showing how assessments impact the educational experiences of teachers and students.

Cholis and Rizqi (2018) used a descriptive quantitative approach to explore the washback of the university entrance examination in Indonesia on English teachers' attitudes and teaching methods. They found that the teachers felt pressured to teach according to the test format, focusing heavily on lesson preparation to prepare the senior high school students for the examination. Meanwhile, the research investigated teachers' perceptions of the washback effect of the SBMPTN. It does not include data from students, who are the primary stakeholders in high-stakes testing. Including student perspectives would provide a more holistic view of how the test impacts teaching and learning processes. Cholis and Rizqi (2018) used a quantitative approach with a Likert scale questionnaire. Meanwhile, this provides measurable insights, qualitative data (such as in-depth interviews) would offer richer, more nuanced insights into how teachers experience the washback effect in different classroom contexts.

The COVID-19 pandemic forced people, including teachers and students, to adjust the learning process so they can foster positive learning outcomes. To prevent the spread of the virus, the Indonesian government implemented a large-scale social restriction (PSBB) program

and closed schools, replacing face-to-face learning with distance learning (PJJ). This spurred a significant shift to online learning, requiring teachers to master digital technologies and organise virtual classes using Zoom, Google Classroom, and WhatsApp. Teachers in Indonesia, particularly those in locations with a restricted Internet connection, confront obstacles in creating dynamic and engaging learning resources. They must also create innovative ways to assess student learning outcomes, such as through project-based assignments, online assessments, and online quizzes. Positive washback was found by [Umam and Zabidi \(2021\)](#), who investigated the impact of online language assessment on the student learning process. Their interviews with 20 English education students suggested that the online assessment improved their understanding of course material and enhanced memory retention, critical thinking, and language proficiency. The students also reported gaining non-linguistic skills, including responsibility, discipline, and time management. However, the participants contested the integrity of the online assessment.

Creating a positive washback effect: Assessment for learning

The preceding discussion demonstrated how national examinations can have a negative impact on classroom teaching practice. Teachers in Indonesia face various challenges from stakeholders, such as parents and students. It should not happen, however, if the school or government takes steps to assist teachers in creating a positive washback effect in the classroom. A ghastly race between schools to see who can get the highest score on the national exam will only put more pressure on teachers. [Harlen and Crick \(2003\)](#) argued that in high-stakes educational contexts, teachers often adopt a more rigid, knowledge-focused teaching style, known as transmission teaching. This approach tends to benefit students who thrive in traditional, lecture-based environments but disadvantages those who prefer more dynamic, hands-on, or creative learning experiences. As a result, it can lower the self-esteem and engagement of these students.

Similarly, [Barksdale-Ladd and Thomas \(2000\)](#) found that teachers perceive standardized testing as detrimental to their teaching effectiveness and students' well-being. The pressure to "teach to the test" limits instructional creativity and places undue stress and anxiety on students. Despite recognizing these negative effects, teachers often feel compelled to focus on test preparation because of the high-stakes nature of the exams. This indicates that the emphasis on standardized testing forces educators to prioritize test performance over meaningful learning, which can undermine both instructional quality and student development. While efficient for exam preparation, the transmission model of teaching fails to accommodate diverse learning styles and stifles opportunities for critical thinking, problem-solving, and creativity in the classroom. Ultimately, the system pressures teachers into practices that may conflict with their professional beliefs and pedagogical goals, creating a tension between meeting external expectations and fostering genuine learning experiences.

Assessment should not be used to narrow the curriculum or distract teachers from teaching to meet professional obligations. Assessment should be used to teach the students rather than to the test. "...while high-stakes accountability testing can provide short-term benefits due to its narrow measures and emphasis on rapid improvement, it quickly degrades and becomes counterproductive" ([Stobart, 2008](#)).

Assessment is designed to generate data to enhance students' learning motivation. This can be achieved by offering students constructive feedback derived from assessment results. Such feedback enables students to understand their strengths and areas for improvement, thereby guiding their learning process. According to Black et al. (2003), assessment can also be utilized to "modify the teaching and learning activities in which they are involved," it provides valuable insights for adapting instructional methods to meet students' needs better and improve.

Implementing the Assessment for Learning (AfL) concept might foster a supportive and motivating classroom environment. AfL emphasizes using assessment not merely as a tool for grading but as an integral part of the learning process. Teachers can create a dynamic and responsive learning atmosphere by regularly assessing and responding to students' progress (Flórez & Sammons, 2013). This approach encourages ongoing dialogue between teachers and students, supports differentiated instruction, and helps build a classroom culture where feedback is viewed as a constructive element of learning rather than a judgment. Ultimately, by prioritizing AfL, educators can enhance student engagement, promote self-directed learning, and improve overall educational achievement.

According to Stobart (2008), assessment for learning means seeking and interpreting evidence for use by learners and their teachers to identify where the learners are in their learning, where they need to go, and how best to get there. This approach was developed in Scotland in response to the need for a more reliable assessment of current pedagogical issues than the conventional one (Flórez & Sammons, 2013). Since the previous activity emphasises a teacher-centred or test-centred approach in the classroom, the idea is to reform the system. Policymakers in Scotland recognised that practitioners and schools are primarily accountable for learners and their families, whose participation is critical to the effectiveness of assessment to support and promote learning (Hutchinson & Young, 2011). The appeal of using this concept (AfL) has also grown in the Asia-Pacific region, including Australia and Hong Kong.

According to Klenowski (2011), implementing the Assessment for Learning (AfL) approach in Queensland, Australia significantly improves the traditional single high-stakes examination. Before the adoption of AfL, student achievement was primarily assessed through a singular examination that served as the sole determinant of academic performance. This method, while straightforward, had notable limitations, as it provided a limited snapshot of a student's abilities and often failed to capture the full spectrum of their learning progress.

Klenowski (2011) argued that AfL offers a more equitable and comprehensive assessment framework. Unlike traditional examinations that assess students based on their performance on a single day, AfL involves ongoing assessment that integrates multiple sources of evidence about a student's learning. This continuous and formative approach allows a more nuanced understanding of a student's abilities, skills, and developmental progress. By focusing on the learning process rather than just the result, AfL provides a richer and more accurate portrayal of a student's educational journey.

AfL identifies and evaluates specific skills and competencies rather than relying on a single performance measure. According to Klenowski (2011), this approach is "more accurate and reflective of an individual's learning and development by identifying the skills and abilities being examined." This means that AfL can better capture the diverse aspects of student

learning and development, including their strengths, areas for improvement, and overall growth.

In Hong Kong, there is a concerted effort among government officials, researchers, and educational practitioners to enhance the quality of teaching and learning by shifting away from an overreliance on standardized exams and embracing the Assessment for Learning (AfL) approach. The region's predominant use of standardized tests has adversely affected educational quality. Berry (2008) noted that if the stakes for success are high, teaching will focus solely on assisting students in passing exams and assisting schools in establishing a good reputation. This high-stakes environment often narrowly focuses on test preparation rather than fostering comprehensive learning experiences. The emphasis on standardized testing can lead to a range of adverse outcomes, such as teaching to the test, which limits the scope of instruction and reduces the opportunity for students to engage in a broad and enriching curriculum. This narrow focus can undermine educational quality by prioritizing test performance over deeper learning objectives and critical thinking skills.

In response to these challenges, there has been a deliberate movement towards reducing the number of high-stakes exams and integrating AfL practices within the educational system. AfL emphasizes using assessment as a formative tool that supports ongoing learning and development rather than merely evaluating student performance simultaneously. Given that Assessment for Learning has been widely implemented at schools in some countries, it will be interesting to discover how teachers in Indonesia perceive the possibility of implementing AfL in classroom practices. Changing teachers' beliefs about reforming the assessment system will be challenging because the government has continuously used the national examination since Indonesia declared independence in 1945. It may not be an uncomplicated way to get through, especially in a country where accountability is based on test scores. According to Stobart (2008), only the most confident teachers and schools will risk encouraging self-regulated learning and self and peer assessment in such assessment cultures. Furthermore, Hutchinson and Hayward (2005) emphasised that those who are honest and willing to change will provide significant improvement.

Conclusion

While high-stakes examinations help monitor educational quality, they frequently limit the curriculum and concentrate disproportionate focus on test preparation, causing increasing stress among students and teachers, moreover, eliminating of the national test in 2016 marks a crucial moment in Indonesian education, providing opportunities to explore more holistic and formative assessment approaches that promote deeper learning and skill development.

The prospect of implementing Assessment for Learning in Indonesia is of great interest in promoting a positive washback effect in classroom practices. The potential of AfL to generate positive washback, particularly in high-stakes testing environments, could provide insights into how assessments can foster deeper learning, enhance student motivation, and improve instructional quality. Future studies should explore how AfL can be effectively integrated into exam-oriented classrooms, how it impacts student engagement and learning outcomes, and how it can alleviate the negative washback often associated with traditional high-stakes assessments. Additionally, examining teachers' and students' perceptions of AfL

in various cultural and educational settings will be essential for designing assessment models that support meaningful learning and positive washback effects.

References

- Aprianto, K. (2013). Validity and washback of English tests in the national examination. *English Education Journal*, 3(1), 106-112.
- Barksdale-Ladd, M. A., & Thomas, K. F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. *Journal of Teacher Education*, 51(5), 384-397.
- Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge journal of education*, 41(3), 287-302.
- Berry, R. (2008). *Assessment for learning*. Hong Kong: Hong Kong University Press
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance exam. *JALT Journal*, 10(1), 15-42.
- Cholis, H. W. N., & Rizqi, F. (2018). Senior high school English teachers' perceptions on a high-stakes test (SBMPTN): A washback study. *International Journal of Education and Literacy Studies*, 6(3), 47-52.
- Dong, M., Fan, J., & Xu, J. (2023). 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*, 43(1), 252-269.
- Estaji, M., & Alikhani, H. (2020). The first certificate in English textbook washback effect: A comparison of teachers' and learners' perceptions. *Teaching English Language*, 14(1), 119-152.
- Fernández, M. P., & Martínez, J. F. (2022). *Evaluating teacher performance and teaching effectiveness: Conceptual and methodological considerations*. Cham: Springer International Publishing.
- Flórez, M. T., & Sammons, P. (2013). *Assessment for learning: Effects and impact*. England: CfBT Education Trust.
- Goodson, I. (2003). *Professional knowledge, professional lives*. McGraw-Hill Education (UK).
- Green, A. (2013). Washback in language assessment. *International Journal of English Studies*, 13(2), 39-51.
- Harlen, W., & Crick, R.D. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice*, 10(2), 169-207.
- Hutchinson, C. & Hayward, L. (2005). The journey so far: Assessment for learning in Scotland. *The Curriculum Journal* 16(2), 225-248.
- Hutchinson, C. & Young, M. (2011). Assessment for learning in the accountability era: Empirical evidence from Scotland. *Studies in Educational Evaluation* 37(1), 62-70.
- Jeffrey, B., & Geoff, T. (2014). *Performativity in UK education: Ethnographic cases of its effects, agency and reconstructions*. BookBaby.
- Klenowski, V. (2011). Assessment for learning in the accountability era: Queensland, Australia. *Studies in Educational Evaluation*, 37(1), 78-83.
- Lichtman, M. (2012). *Qualitative research in education: A user's guide*. Thousand Oaks, CA: SAGE Publications.

- Madaus, G. F. (1988). The distortion of teaching and testing: High-stakes testing and instruction. *Peabody Journal of Education* 65(3), 29–46.
- Mons, N. (2009). *Theoretical and real effects of standardised assessment*. Brussel: Eurydice.
- Pitoyo, M. D., & Asib, A. (2020). Gamification-based assessment: The washback effect of Quizizz on students' learning in higher education. *International Journal of Language Education*, 4(1), 1-10.
- Puspitasari, M. (2020). *Investigating the washback effect of the national examination on Indonesian practices: Perceptions of teachers, students and parents of test impact*, [Doctoral dissertation, University of Glasgow].
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5-29.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Oxon: Routledge.
- Sukyadi, D., & Mardiani, R. (2011). The washback effect of the English national examination (ENE) on English teachers' classroom teaching and students' learning. *K@ Ta*, 13(1), 96-111.
- Tsagari, D. (2011). Washback of a high-stakes English exam on teachers' perceptions and practices. *Selected Papers on Theoretical and Applied Linguistics*, 19, 431-442.
- Umam, A., & Zabidi, O. W. (2021). The washback effect in online language assessment on student's learning interest. *Pioneer: Journal of Language and Literature*, 13(2), 162-172.
-

Biographical Note

MAYA PUSPITASARI, PhD. is working at Universitas Terbuka, Indonesia.