
The washback of the abolition of the national examination in Indonesia

Ifan Rikhza Auladi

English Education Department, Faculty of Languages and Communication Science, Universitas Islam Sultan Agung, Indonesia

MEd Teaching English to Speakers of Other Languages (TESOL), The University of Exeter, United Kingdom

***Corresponding Author**

Email: ifanrikhza@unissula.ac.id

How to cite (APA 7th style): Auladi, I. R. (2024). The washback of the abolition of the national examination in Indonesia. *Indonesian Journal of Education and Pedagogy*, 1 (1), 47-62. <https://doi.org/10.61251/ijoe.v1i1.50>

Abstract

The National Examination (NE) is abolished in 2021, marking the beginning of a new educational evaluation paradigm in Indonesia. The regulation considers some reasons: the suitability of the test, the development of teaching methods, and the enhancement of the quality of national education. Under the abolition of NE, the regulation constitutes the establishment of School-Based Assessment (SBA), conducted by the teachers in the forms of portfolios, assignments, tests, and/or other assessments following the regulation from each school. Therefore, the present research aims at investigating English teachers' perception towards the abolition of NE and the washback of the abolition of NE on the teaching plan of English language learning in the aspect of classroom instructional behaviours and the content of the curriculum. The sequential mixed-methods research (MMR) works with quantitative and qualitative data from 51 respondents in an online questionnaire and four teachers participated in a semi-structured interview. It was then found that the establishment of SBA replacing NE obtains positive responses from the teachers as it is perceived more reliable with objectives of the English curriculum, which attempts to elevate students' high order thinking skills. The abolition of NE triggers the teachers to vary their teaching methods and techniques by using more communicative approaches; however, the change in assessment does not influence teachers' domination in classroom activities and talking activity. The teachers highly emphasize the use of previous exam papers in the preparation of SBA. In addition, the teachers remain highly prioritizing students' success in SBA, which retains exam papers, as the level of students' language proficiency is relatively low. Referring to the findings, the related parties can measure the washback of the abolition of NE for further improvement in the language assessment.

Keywords: National Examination; School-Based Assessment

INTRODUCTION

The abolition of the National Examination (NE), officially conducted in 2021, is a starting point for a new educational assessment paradigm in Indonesia. The new assessment system is a response to numerous studies about the washback of the establishment of NE, which is

typically associated with exam papers as high-stakes tests. Most studies argue that NE has caused problematic issues concerning the low motivation in teaching and learning and the retainment of teaching behaviours (Saukah & Cahyono, 2015; Sutari, 2017).

Under the 2021 assessment system, School-Based Assessment (SBA) is applied to determine students' graduation based on their involvement and attainment in learning process (Kemdikbud, 2021a). SBA emphasises the process of teaching and learning rather than the result of the tests. The change of the assessment from NE to SBA is intended to elevate the quality of national education (Kemdikbud, 2021b).

In response to the abolition of NE and the establishment of SBA, teachers have important roles in the transition process of new assessment system as the national curriculum promotes independent teaching and learning. The teachers hold the responsibility to organise classroom activities and assess the students based on their abilities and resources; therefore, it is essential to understand teachers' perceptions towards the abolition of NE and the establishment of SBA.

The stakeholders should reflect from the establishment of School Exam and low-stakes tests in the previous eras. The establishment of School Exam in the 1970s caused the low quality of most graduates as the impact of the regulation, which handed over the authority in assessment from the government to the schools (Kemdiknas, 2010; Nugroho, 2009). Within the next era, the implementation of NE as a low-stakes test in 2015 confirmed that the teachers still perceived NE as a high-stakes test, indicated by the use of the previous examination for students' exercises (Sutari, 2017).

Considering research background, the reason for choosing the topic is a fierce debate on agreement and disagreement of NE to measure students' academic achievement as a requirement for the graduation. Many researchers argue that the washback of NE has adversely affected classroom teaching and narrow curriculum aims and objectives. The regulation on the abolition of NE and the establishment of SBA should consider the possibility to repeat the problems of the previous assessments. Therefore, the research related to teachers' perceptions, classroom instructional behaviours, and the curriculum contents is important to assure the education quality following the change of testing and assessment.

Washback in language testing and assessment

The nature of washback

The linguistic terms washback and backwash are frequently interchangeable in language testing (Alderson, 2004; Hughes & Hughes, 2020). In the present research, the term washback, as the most prominent in applied linguistics (Cheng & Curtis, 2004), refers to both washback and backwash. Backwash is more prevalent in general education literature than washback, most frequent in language education, specifically in "British applied linguistics" (Andrews, 1994; Alderson & Wall, 1993). Following Cheng (2005), the term washback or backwash is retained in the direct quotations from the authors.

Hughes and Hughes (2020, p. 3) define backwash as "the impact of testing on teaching and learning". Messick (1996, p. 241) affirms that "washback ... refers to the extent to which the introduction and use of a new test influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning". Cheng (2005, p. 112) specifies washback as "an intended or unintended (accidental) direction and function of curriculum change on aspects of teaching and learning by means of a change of

public examinations". Numerous definitions outline the relation of a test with the impact on aspects of teaching-learning, which cannot be separated in the discussion of educational assessment.

"Backwash" has been evolving to cover the impact of the test on learners and teachers as well as general educational systems and large society (Hughes & Hughes, 2020). Bachman and Palmer (1996) similarly state that washback affects individuals and educational and social systems. For instance, the establishment of high-stakes tests is intended to correlate and rate people, institutions, or national systems (Chapman & Snyder, 2000). The difference between high and low-stakes tests also creates the discrepancy in the washback of the test on teaching and learning process and the perception from the teachers, learners, and bureaucrats (Shohamy et al., 1996). Shohamy et al. (1996) explain that high-stakes tests influence teaching activities, time allotment, course generation, test awareness, and status-matter, while low-stakes tests generate less effect in the same aspects.

Hung (2012) notes that washback demonstrates positive and negative influences of the tests on teaching and learning. Washback can be positive or negative depending on whether it helps or hinders learners and/or programme personnel achieve their educational objectives (Bailey, 1996). In addition, the attainment of the assessment purpose and the achievement of the curriculum aims and objectives indicate positive or negative washback from the test (Brown & Hudson, 1998). Normally, the audience has significant role in distinguishing between positive and negative washback (Watanabe, 2004). Therefore, the washback of the test to the target person varies as the impact can also be different depending on experiences and beliefs (Xu & Liu, 2018).

Washback perceived success when showing the impact on the process of teaching and learning, the effectiveness of format tests, and the influence of test attainment (Fulcher, 2010). In the attainment of the purpose, Cheng and Curtis (2004) state that the test has a positive impact if the teachers and learners are passionate to achieve the purposes. An increase in particular movement or enthusiasm indicates the existence of positive washback from the test (Alderson & Wall, 1993).

Brown and Abeywickrama (2010) point out that the test becomes meaningful when it positively influences teachers' methods and materials, and provides the students adequate preparation, feedback, and chance to elevate language skills and gain best achievement. For example: the positive washback of the test for the teachers: using more authentic materials from the news, establishing meaningful learning activities, encouraging more students' participation, and using an integrated approach in teaching (Lam, 1995). The use of authentic materials tends to relate to process rather than test-based learning.

Despite its positive impact, a test may negatively influence the teaching and learning. Hughes and Hughes (2020) set a fundamental understanding of test-related discussion, in which the easiest test tends to be conducted rather than the most important test. Shohamy (2001, pp. 4&47) explains that "the test takers need to match their performances to the tests rather than the tests to the test takers" as the consequences of the test result, which has strong influences in revising curriculum. Cheng and Curtis (2004) confirm that tests, as the primary tools, can change the education system without considering other factors.

Wall (2005) explains that an examination has consequences on teachers' attitude, teaching content and methodology, and students' assessment. Wall reveals that many teachers indicated that their goal in teaching was to assist the students pass the

examination, which meant giving greater attention to the areas that were assessed (Wall, 2005). Curriculum distortion is also perceived as the impact of examinations, indicated by the ignorance of untested subjects and activities and the overwhelming exam preparation (Vernon, 1956 in Alderson & Wall, 1993). Green (2014) further reveals that teachers' focus on teaching materials may change as they prioritise the assessments than the curriculum.

Theoretical Framework for Washback Studies

Alderson and Wall (1993) outline fifteen washback hypotheses, as follow:

- 1-A test will influence teaching.
- 2-A test will influence learning.
- 3-A test will influence **what** teachers teach.
- 4-A test will influence **how** teachers teach.
- 5-A test will influence **what** learners learn.
- 6-A test will influence **how** learners learn.
- 7-A test will influence the **rate** and **sequence** of teaching.
- 8-A test will influence the **rate** and **sequence** of learning.
- 9-A test will influence the **degree** and **depth** of teaching.
- 10-A test will influence the **degree** and **depth** of learning.
- 11-A test will influence attitudes to the content, method, etc. of teaching and learning.
- 12-Tests that have important consequences will have washback.
- 13-Tests that do not have important consequences will have no washback.
- 14-Tests will have washback on **all** learners and teachers.
- 15-Tests will have washback effects for **some** learners and **some** teachers, but **not** for others (Alderson & Wall, 1993, pp. 120-121).

Alderson and Hamp-Lyons (1996, p. 282) clarify the Fifteen Washback Hypotheses that "Alderson and Wall's intention is not to argue for or against any one of these hypotheses, but to 'lay out the territory' so that the questions that studies of washback must address can more clearly be seen". Xu and Liu (2018) further classify the Fifteen Washback Hypotheses from Alderson and Wall to differentiate the distinction between the areas of teaching (1, 3, 4, 7, 9, and 11) and learning (2, 5, 6, 8, 10, and 11), content (3 and 5) and method (4 and 6), practice (3–10) and attitude (11), and tests and participants (12–15). The present research employs seven hypotheses from Alderson and Wall in the areas of teaching (1, 3, 4) and tests and participants (12–15) considering the focus of the discussion.

The second framework is Hughes' Washback Trichotomy, which constitutes the differences of certain aspects, such as participants, process, and product in teaching and learning. Hughes (1993, as cited in Bailey, 1999, p. 9) proposes that "to clarify our thinking on backwash, it is helpful, ... to distinguish between participants, process and product in teaching and learning, recognizing that all three may be affected by the nature of a test".

Onaiba (2013, p. 52) further explains Hughes's Washback Trichotomy, as follows:

Participants are ... "all of whose perceptions and attitudes towards their work may be affected by a test" (cited in Bailey, 1999: 9). The constituent *process* refers to any action or interaction taken by the participants that may (or may not) contribute to learning Finally, *product* covers what is achieved.

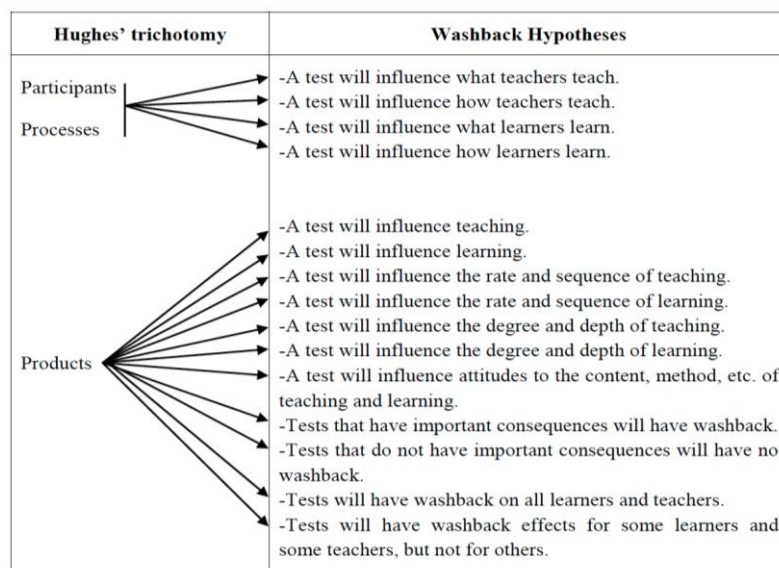


Figure 1. Hughes's Washback Trichotomy
(Onaiba, 2013, p. 53)

Washback on Teachers' Perceptions, Classroom Instructional Behaviours, and Curriculum

In the washback on teachers' perceptions, Smith (1991) explains that testing affects teachers' feelings, such as the unpleasant feelings upon the publication of test score, the dissonant and alienated feelings upon the quality of the test and the pressure of the score target, and the anxious and guilt feelings of students' failure. Cheng (2005) finds that teachers' perception towards the change of the HKCEE shows implicit agreement with the fundamental purposes of the new examination, although the teachers feel anxious upon students' low ability and limited preparation to practice for the examination.

In the washback on classroom instructional behaviours, multiple-choice testing reduces teachers' creativity in teaching methods (Smith, 1991). McNamara (2000) gives an example on the washback of testing as "multiple choice format item tests of grammar or vocabulary knowledge may inhibit communicative approaches to learning and teaching". Consequently, the teachers mainly dominate classroom activities, which significantly takes students' chance to practice their English (Xu & Liu, 2018). In addition, teachers' propensity to teach to the test is frequently noted as a barrier to implement innovative instructional approaches (Chapman & Snyder, 2000).

Meanwhile, Cheng (2005) and Wall (2005) confirm teachers' ability and commitment in changing teaching methods as a problematic issue in the effect of testing. Cheng (2005) argues that the change in the test may generate the teachers to develop various classroom activities; however, the teachers still retain their primary views and thoughts in the establishment of teaching and learning process. The teachers have changed their teaching based on the new test although they retain common teaching style indicated by their domination in classroom activities (Cheng, 2005).

Wall (2005) states that teachers' capability and commitment are the most difficult to implement new teaching methods because of limited training and less confidence. He further explains that the impact of examinations cannot alter teaching practices if the

teachers do not perform their commitment to implement innovative notions and willingness to elevate abilities in the experiment, assessment, and modification to the new methods. The teachers also need supporting environment to change their teaching methods in classroom activities, schools, local systems, and higher levels of educational administration (Wall, 2005).

Morris (1972 in Alderson & Wall, 1993) states that examination has essential roles in establishing the curriculum. In line with Morris, Orafi and Borg (2009) argue that examination becomes the most influential factor in interpreting and implementing the curriculum. Saif (2006) states that the test immediately affects the teaching substance as the test result determines the change of content and curriculum in language teaching (Wall & Alderson, 1993; Cheng, 1997).

In the content of materials, testing reduces teachers' time availability required for instruction and focuses on the uncovered materials in the test (Smith, 1991). Consequently, the test decreases teachers' capacity to adapt, invent, or deviate and narrow the conceivable curriculum (Smith, 1991). Smith et al. (1991) confirm teachers' tendency not to prioritise the materials that the external test does not cover due to the pressure from other parties to succeed in the test.

Another washback on teaching materials is presented by Andrews (1994), who finds out that two-thirds of class time is spent working with published exam-focused materials, with practise tests taking up the greatest time. In line with Andrews' finding, Green (2014) confirms that teachers and students naturally focus on the information, skills, or talents that they perceive the exam covers and pay less attention to those that do not appear to be covered when preparing for an assessment. The teachers tend to specify their teaching content on what can be predicted from the scope of the test, which is highly predictable (Hughes & Hughes, 2020). Wall (2005) reveals teachers' reluctance to use integrated-skill material equally in the teaching and learning process because the examination does not test four-language skills.

METHOD

Research Design

As a way of looking at the world (Mertens, 2010), the paradigm of the present research is the pragmatic paradigm employing sequential mixed-methods research (MMR). MMR connects two contrastive philosophies between positivism and constructivism (Johnson & Onwuegbuzie, 2004). It is "working primarily within the pragmatist paradigm and interested in both narrative and numeric data and their analyses" (Teddlie & Tashakkori, 2009, p. 12).

The sequential MMR of the present research employs the quantitative and qualitative data to investigate the washback of the abolition of NE in Indonesia towards teachers' perception and the teaching plan on classroom instructional behaviours and the content of the curriculum. The quantitative approach is related with the philosophy of positivism, while "qualitative methodologists advocate a more 'artistic' approach to research, adhering to other worldviews (such as postmodernism)" (Fraenkel & Wallen, 2009, p. 559).

As the washback of the test is considered a complex phenomenon (Alderson & Wall, 1993), using several different methods is recommended to gather data (Wall, 2005). Consequently, the use of MMR is preferred to provide a comprehensive understanding of the washback of the revised exam on teaching and learning (Cheng, 2005). MMR offers some

advantages, such as a more comprehensive knowledge of research issues than either strategy alone (Fraenkel & Wallen, 2009), additional value to the use of each method on its own, more evidence when both types of data are obtained, and further understandings arise when both types are included (Creswell & Zhou, 2016).

Data Collection

The present study research was conducted within two phases of data collection methods, called the explanatory sequential design (Creswell & Clark, 2017). The first phase used close-ended questionnaires to measure the general perceptions from the teachers. The second phase customized individual interviews as the foundation to investigate the result of the first phase in more profound understanding and validate the result of quantitative data (Creswell & Clark, 2017). Both phases were sequentially analysed and interpreted to determine the findings and conclusion of the research.

In the first phase, the instrument of quantitative data collection employed internet-based questionnaires, in which individuals took part in a survey meant to be completed and submitted online (Rea & Parker, 2014). The questionnaire was created by using Google Form and distributed to English teachers of Senior Secondary Schools in early October 2021. The questionnaire was adapted from Onaiba's model (2013), which was perceived as reliable with the present research, as Onaiba investigated the washback of the revised examination from teachers' perspective, instructional practices, and the content of the curriculum. The second phase employed a semi-structured interview, which provides flexibility and the pursuit of unanticipated lines of inquiry throughout the interview (Grix, 2004). As the interview allowed the researcher to ask the questionnaire results (Dörnyei & Taguchi's, 2010), it provided a broader explanation from the quantitative data (Ziegler & Kang, 2016). Therefore, the interview aimed to seek deep understandings from the questionnaire results related to the abolition of NE.

A quota sampling method was used to divide the participants into subgroups based on the grade levels they were teaching. Four participants from the previous phase participated in the individual interview, which was delivered in Indonesian. The researcher considered participants' responses and backgrounds to select interviewees in the second phase. Before the interview was conducted, the researcher distributed interview guidelines to the participants in order to understand the research purposes clearly for their convenience. As the interview was preferred to use a medium matched with the interviewee's availability (Brinkmann, 2014), the interview was conducted using a Zoom application, a common online meeting tool. The interview was recorded with the permission from the interviewees as the use of video increased the impact of the interview by providing nonverbal communication as well as the physical surroundings of the interview (Gilham, 2000). At the end, the transcription result was translated into English before being analysed.

Participants

The present explanatory sequential research conducted a quantitative and qualitative sampling strategy in selecting the participants within two phases of data collection methods. The selection of the sample, as information resources, for quantitative data collection considered the convenience of the researcher and respondents (Fraenkel & Wallen, 2009). The researcher chose the participants based on their willingness and availability to provide the data, assist understandings (Creswell, 2015), and contribute appropriate knowledges of the topic area (David & Sutton, 2011). The selected participants

were then asked to establish snowball sampling to obtain more respondents (Auerbach & Silverstein, 2003).

In the first phase, the data was derived from 51 respondents who voluntarily contributed to the questionnaire. The questionnaire was previously distributed to the English teachers considering the convenience, willingness, and availability from the targeted respondents. After selecting English teachers from his hometown, Semarang, Indonesia, the researcher requested the respondents to share the questionnaire to other colleagues (the snowball sampling). The respondents in the first phase were dominated by female teachers; that was 66.7% of the total sample. The sample showed that senior teachers were the majority comprising 54.9% of the respondents. The sample also displayed 30 respondents (58.8%) teaching Grade 12 and 21 teachers (41.2%) teaching Grade 10 and/or 11.

In the second phase, a quota sampling method was applied to select four interviewees based on their questionnaire responses. The teachers of Grade 12 were prioritised to participate in the interview as they held responsibilities to teach the students preparing for SBA. The profiles of the interviewees were two teachers of all grades (Grade 12, 11, and 10) and two teachers of two grades (Grade 12 & 11 and Grade 11 & 10) in Senior Secondary School. Two participants were considered as senior teachers by having more than 15 years of teaching experience, while two others had less than 15 years.

Data Analysis

The research employed quantitative and qualitative data analysis to answer the research questions. Firstly, in the implementation of quantitative data analysis, the data file was tabulated by Statistical Package for the Social Sciences (SPSS) for a more robust analysis (Rea & Parker, 2014). The mean and standard deviation scores became parts of descriptive statistics from the analysis of the questionnaire results (Pallant, 2010). Both mean and standard deviation provide insights for overall comparison among items or groups (Brown, 2014). The responses in a five-point Likert scale of agreement were valued as ordinal data, where 1 = 'Strongly Disagree', 2 = 'Disagree', 3 = 'Undecided', 4 = 'Agree', and 5 = 'Strongly Disagree'. Similarly, other responses were valued in a 5-point Likert scale of frequency, where 1 = 'Never', 2 = 'Seldom', 3 = 'Sometimes', 4 = 'Often', and 5 = 'Always'. Nominal data was applied for the dichotomous and multiple choice. The answer of more than one in the multiple-choice was counted as an individual dichotomous question, with the value 1 for the answer.

Secondly, in the process of the qualitative data analysis, the data from all parts of the corpus were taken and organised into thematic topics (Holliday, 2002). The process of coding, which comprises of selecting, defining, comparing, and grouping, was used to organise transcripts and discover patterns in organisational structures (Auerbach & Silverstein, 2003). The codes with similar ideas were grouped using a NVivo application based on the themes, which the researcher interpreted afterwards.

At the end of the analysis process, the data from interviews and questionnaires were analysed to answer the research questions. The result was then presented and discussed in the findings and discussion, which eventually yielded the research conclusions. Additionally, some recommendations were proposed for further researches.

DISCUSSION

The questionnaire result showed that teachers' perception towards the abolition of NE indicated the agreement upon the abolition of NE to prepare the students for their future education. The result is in line with the national education goal emphasising higher-level thinking and more practical skills (Kemdikbud, 2021b). The result also reveals teachers' concern on the leak issue in the implementation of NE. Teachers' awareness on the leak issue confirming the study about the pressure to the teachers upon students' success in the exam (Smith, 1991).

However, there was a different argument in the interview as the abolition of NE perceived by the teacher driving English teaching and testing to a different standard. The assessment was practically conducted by the teachers based on their ability and facility. Therefore, the standard of learning and assessment is considered useful to guide the teachers in delivering materials and assessments as the consequence of the abolition of NE.

When being asked about the perception of the key changes in the use of SBA compared to NE, the teachers showed their agreement in the implementation of SBA related to the principles of the current English curriculum. The interview also showed the advantage of establishing SBA as the assessment adjusted to the teachers' material and students' ability. This finding confirms the study from Davison and Leung (2009) stating that assessment correlates with the curriculum content.

The questionnaire result on teachers' belief and perception towards SBA showed that the teachers perceived the establishment of SBA should include other aspects such as speaking, listening, and writing. The result of the interview confirmed teachers' authority to assess the students based on the ability and resources. However, the assessment application could not cover an integration of language skills as some aspects are difficult to apply considering teachers' ability and facility. For example, the need for English-fluent interlocutors or native speakers to read the questions is a barrier to make the questions assessing listening skills. The teachers also use partial parts of the previous examinations, which are likely to be predictable for the students and unreliable with the current materials. The finding echoes Cheng's (2005) study about problematic teachers' ability and commitment to reform teaching methods. Consequently, the teachers would be reluctance to teach integrated language skills if the assessment only tests some particular skills (Wall, 2005).

The result related to additional works or pressures in the establishment of SBA to replace NE demonstrated that the teachers prioritise students' success on the test even though SBA, as a low-stakes test, was conducted by the teachers. The finding echoes Puspitasari's (2020) study confirming that the students' achievement in a low stakes-test highly affects schools' credibility. In this case, the teachers hold two contrastive responsibilities: assuring the assessment quality to fulfil society trust (Davison & Leung, 2009) and considering students' ability to succeed in SBA (Puspitasari, 2020).

In the discussion on the assessed areas of language skills, the teachers agreed on the integration of language skills in the assessment. The finding aligns with teachers' perception of SBA, expecting to include other aspects. However, the implementation of integrated skills in the assessment depending on teachers' ability as revealed in the interview related to teachers' competence to create particular skill questions. The finding aligns with Wall's (2005) study confirming that the lack of ability is the barrier to test language skills.

In line with the previous discussion, teachers' perception towards the importance of studying English confirmed that English is important to enable the students to be competent in language skills for their educational and cultural purposes. Teachers' awareness on the importance of English for students' education would change the paradigm of language learning and assessment being more communicative and integrative. The result supports the finding on the perception towards the abolition of NE, the key changes in the use of SBA compared to NE, the use of SBA to replace NE, and the assessed areas of language skills.

The discussion in teachers' perception showed that students' current proficiency level is the most difficult aspect in language teaching. The learning environment and limited time allotment also cause less exposure for the students to practice their language skills. The finding supports Cheng's (2005) study that students' English proficiency causes teachers' anxiety in the test and Wall's (2005) research that the environment influences teaching methods. However, the finding challenges the argument from Shohamy et al. (1996) confirming that low-stakes tests generate less effect including time allotment.

In the initial discussion of the washback on classroom instructional behaviours, the result of the questionnaire displayed a positive washback in teaching methods and techniques. The use of a more communicative approach, followed by adopting new methods, became the major choices of teachers' change in teaching methods and techniques. From the interview result, the teachers expressed their independent authority in implementing a more communicative approach and adopting new teaching methods, such as practicing debate, which cannot be conducted beforehand. The change of methods and techniques in teaching indicates meaningful washback from the test (Wall, 2005) and confirms that "a test will influence **how** teachers teach" (Alderson & Wall, 1993), such as **methodology** and **methods**, to prepare the test (Lam, 1995). The finding also displays the variation of teaching styles, which did not occur in Sutari's (2017) study investigating the change of NE from a high to low-stakes test.

From the result related to classroom activities, the teachers tended to give more homework to the students. By providing more assignments as expressed by the teacher, the students are expected to learn more as a consequence of the abolition of NE, which drill the students with multiple-choice questions. The finding supports the hypothesis of "a test will influence teaching" (Alderson & Wall, 1993).

The result of teachers' talking activity correlated with the medium of instruction in the classroom. The result shows that the teachers highly tend to talk to the whole class as the indication of teacher-centred learning, causing less exposure for the students to practice language skills. The finding supports the study from Andrews (1995) and Xu and Liu (2018) reporting teachers' domination in talking time and classroom activities as significant concerns in washback studies. In addition, the finding is described by Hughes (1993) as the *process*.

The questionnaire result regarding teachers' recommendations in learning strategies showed that the teachers suggested the students to use previous exam papers similar to the content of SBA. This result contradicts with the significant changes in teaching in the context of SBA to replace NE – TQ8 (using a more communicative approach in teaching). Even though the teachers react positively towards the abolition of NE and the establishment of SBA (TQ 1 – 3), they still emphasise the use of exam papers, covered in SBA, for students' exercise in learning strategies. This finding echoes Green's (2014) study confirming that the teachers would prioritise the preparation for the assessment and the tested materials. From

the interview result, the domination of exam papers in SBA demonstrates the impact of teachers' limited skills and facilities in designing integrated-skill assessments (Chin et al., 2019). In addition, the finding confirms that the easiest test is more preferable than the most important test (Hughes & Hughes, 2020). However, the finding contradicts with the hypothesis from Alderson and Wall (1993) outlining that tests that do not have important consequences will have no washback; SBA, as a low-stakes test, influences the focus of teaching and learning.

The teachers' consistency in the content and format of the classroom tests and SBA indicates that the teachers adjust the content and format of the classroom tests similar with SBA by familiarising the students with the pattern of the questions (see TQ4). The finding correlates to the consistency in teachers' teaching and test across classes in the same grade. On the other hand, the teachers tend to vary their teaching and test across grades due to different target of learning and assessment. As the students may have different English levels across classes (see TQ7 - students' English level as the most challenging teaching), the teachers adapt their teaching and test based on students' ability. The result also confirms that the teaching and test are different across grades, indicating teachers' preparation for the student to face SBA in the final school year.

Following classroom instructional behaviours, the washback of the abolition of NE in the content of the curriculum is presented. In teaching and learning resources, there are two impacts from the 2021 assessment system related to the technology advancement. At first, the teachers retain the classical class using a black or whiteboard to explain the materials. Secondly, the popularity of social media applications triggers internet-based learnings, which can be assessed through personal devices. Both impacts show that students' affordability determines the availability of learning resources, as stated in the interview, which depicts students' ability to reach the current sophisticated learning resources. The teachers adjust their learning resources based on students' ability whether they are prepared to assess internet-based learnings or to retain the classical class. The result supports the hypothesis that "tests will have washback effects for **some** learners and **some** teachers, but **not** for others" (Alderson & Wall, 1993).

With the focus of language aspects on speaking and pronunciation sections along with reading comprehension lessons, teachers' content of the current teaching syllabus performs two main aspects of SBA. The focus on speaking and pronunciation correlates to the significant changes in teaching methods and techniques using a more communicative approach (TQ8). Then, a reading comprehension lesson links to teachers' recommendation to prepare SBA mainly covering exam papers (TQ9). The finding confirms that "a test will influence the **degree** and **depth** of teaching" (Alderson & Wall, 1993) classified as the *process* (Hughes, 1993).

In the next point, teachers' arrangement and selection of teaching content and material showed the combination of current textbooks and teachers' materials. The result indicates that the teachers attempt to use standard materials from the national curriculum and use a more communicative approach and create new teaching methods (see TQ8). The teachers positively implement a more communicative approach in teaching by combining current textbooks and teachers' self-materials.

Related to the result confirming the teachers as the most responsible people in designing the teaching materials, the finding supports the discussion on teachers' authority to establish classroom activities. By having authority, the teachers adjust their materials

based on teachers' resources and students' needs. In addition, the teachers show their confidence related the most influencing factor in teaching, which is 'teaching experience'. The result indicates that teachers' experiences generate a higher influence and confidence on teaching styles than other aspects, including the external exam and the expectation from related parties. The finding confirms that the washback of the test varies depending on experiences and beliefs (Xu & Liu, 2018).

CONCLUSION

The discussion from the findings yields the conclusion related to the research questions. Firstly, English teachers perceive that the abolition of the National Examination (NE) positively impacts their English teaching under the curriculum. Conversely, the establishment of School-Based Assessment (SBA), which retains the use of exam papers, to replace NE does not change teachers' perception significantly as their priority aims to enable the students to obtain success in the test.

Secondly, regarding the washback on the teaching plan of English language learning, the abolition of NE generates various impacts in the aspect of classroom instructional behaviours and the content of the curriculum. The teachers implement more communicative approaches in teaching methods and techniques, even though the use of exam papers in SBA is retained. As the use of black/whiteboard and the internet dominates teachers' choice in teaching aids, speaking and pronunciation sections along with reading comprehension lesson become teachers' focus of language aspects from current teaching syllabus. The use of a mixture of current textbooks and teachers' materials in the arrangement and selection of teaching content and material show teachers' attempt to develop their teaching from the curriculum based on the resources and needs.

Implication

The research finds that the abolition of NE and the establishment of SBA, which give an authority to the teachers, does not automatically transform language learning into more communicative in the classroom activities and integrated language skills in testing. There is a tendency for the assessment conducted based on teachers' ability to test the students even though the teachers perceive the importance of integrated language assessment. The research provides initial insights regarding the washback of the abolition of NE conducted in the early period of the 2021 assessment system. The related parties can identify the positive and negative washback of the abolition of NE and the establishment of SBA.

Limitations and directions for further research

The present research has two significant limitations in time and methods. At first, the research is conducted in the early period of the abolition of NE. The change in assessment has not significantly affected the teaching and learning process. However, the limitation will stimulate further studies as practical recommendations can be presented for the improvement of language assessment in Indonesia. Secondly, more instruments and samples should be involved in the research design. For instance, the researchers adopt an observational data in the classroom practice and document analysis of SBA through qualitative approaches. The customized students' questionnaire in a quantitative approach may result stronger arguments in the discussion and conclusion.

ACKNOWLEDGEMENTS

We thank English Study Program, Universitas Islam Sultan Agung, Indonesia and the University of Exeter, the United Kingdom for supporting us to produce scientific publication on reputable journals.

REFERENCES

- Alderson, J. C. (2004). Foreword. In Cheng, L., Watanabe, Y., & Curtis, A. (Eds.), *Washback in Language Testing: Research Context and Methods* (pp. ix-xii). Lawrence Erlbaum Associates.
- Alderson, J. C. & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*. 13(3), 280–297. <https://doi.org/10.1177/026553229601300304>
- Alderson, J. C. & Wall, D. (1993). Does washback exist? *Applied Linguistics*. 14(2), 115-129. <https://doi.org/10.1093/applin/14.2.115>
- Andrews, S. (1994). Washback or washout? The relationship between examination reform and curriculum innovation. In D. Nunan, R. Berry and V. Berry (Ed.), *Bringing about change in language education* (pp. 67-81). University of Hong Kong.
- Auerbach, C. F. & Silverstein, L. B. (2003). *Qualitative data: An Introduction to Coding and Analysis*. New York University Press.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*. 13(3), 257–279. <https://doi.org/10.1177/026553229601300303>
- Brinkmann, S. (2014). Unstructured and Semi-Structured Interviewing. In Leavy, P. (Ed.), *The Oxford Handbook of Qualitative Research* (pp. 277-299). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199811755.013.030>
- Brown, H. D. & Abeywickrama, P. (2010). *Language Assessment: Principles and Classroom Practices*. Pearson Longman.
- Brown, J. D. (2014). *Mixed Methods Research for TESOL*. Edinburgh University Press.
- Brown, J. D. & Hudson, T. (1998). The Alternative in Language Assessment. *TESOL Quarterly*. 32(4), 653–675. <https://doi.org/10.2307/3587999>
- Chapman, D. W. & Snyder Jr, C. W. (2000). Can high stakes national testing improve instruction: reexamining conventional wisdom. *International Journal of Educational Development*. 20(6), 457-474. [https://doi.org/10.1016/S0738-0593\(00\)00020-1](https://doi.org/10.1016/S0738-0593(00)00020-1)
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*. 11, 38–54. <https://doi.org/10.1080/09500789708666717>
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.
- Cheng, L. & Curtis, A. (2004). Washback or washout: A review of the impact of testing on teaching and learning. In Cheng, L. Watanabe, Y., and Curtis, A. (Eds.), *Washback in Language Testing: Research Context and Methods* (pp. 3-17). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781410609731>
- Chin, H., Thien, L. M., & Chiew, C. M. (2019). The Reforms of National Assessments in Malaysian Education System. *Journal of Nusantara Studies*. 4(1), 93-111. <https://doi.org/10.24200/jonus.vol4iss1pp93-111>
- Creswell, J. W. (2015). *Educational Research Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (5th ed.). Pearson.
- Creswell, J. W. & Clark, V. L. P. (2017). *Designing and Conducting Mixed Methods Research* (3rd ed.). Sage Publications.
- Creswell, J. W. & Zhou, Y. What is mixed methods research? In Moeller, A. J., Creswell, J. W., & Saville, N. (Eds.), *Second Language Assessment and Mixed Methods Research* (pp. 35-50). Cambridge University Press.
-

-
- David, M. & Sutton, C. D. (2011). *Social Research: An Introduction*. Sage Publications.
- Davison, C. & Leung, C. (2009). Current Issues in English Language Teacher-Based Assessment. *TESOL Quarterly*. 43(3), 393-415. <https://doi.org/10.1002/j.1545-7249.2009.tb00242.x>
- Dörnyei, Z. & Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203864739>
- Fraenkel, J. R. & Wallen, N. E. (2009). *How to Design and Evaluate Research in Education* (7th ed.). McGraw-Hill Higher Education.
- Fulcher, G. (2010). *Practical Language Testing*. Hodder Education.
- Gillham, B. (2000). *The Research Interview*. Continuum.
- Green, A. (2014). *Exploring Language Assessment and Testing*. Routledge. <https://doi.org/10.4324/9781315889627>
- Grix, J. (2004). *The Foundations of Research*. Palgrave Macmillan.
- Holliday, A. (2002). *Doing and Writing Qualitative Research*. Sage.
- Hughes, A. & Hughes, J. (2020). *Testing for Language Teachers* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781009024723>
- Hung, T.A. (2012). A washback study on e-portfolio assessment in an English as a foreign language teacher preparation program. *Computed Assisted*. 25(1), 21-36. <https://doi.org/10.1080/09588221.2010.551756>
- Johnson, R. B. & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*. 33(7), 14-26. <https://doi.org/10.3102/0013189X033007014>
- Kemdikbud. (2021a). *Circular Letter of Minister of Education and Culture Number 1 Year 2021*. Ministry of Education and Culture.
- Kemdikbud. (2021b). *Daftar Tanya Jawab Kebijakan Ujian Nasional (UN) [The List of Questions and Answers on the Policy of the National Examination]*. <https://www.kemdikbud.go.id/main/tanya-jawab/tanya-jawab-ujian-nasional>
- Kemdiknas. (2010). *Sejarah Ujian Nasional [History of the National Examination]*, retrieved from <http://www.kemdiknas.go.id/orang-tua/ujian-nasional/sejarah-ujian-nasional.aspx>.
- Lam, H. P. (1995). Methodology washback - An insider's view. In Nunan, D., Berry, R., & Berry, V. (Eds.), *Bringing about change in language education* (pp. 83-102). The University of Hong Kong.
- Lune, H. & Berg, B. L. (2017). *Qualitative Research Methods for the Social Sciences* (9th ed.). Pearson.
- McNamara, T. (2000). *Language Testing*, Oxford University Press.
- Mertens, D. M. (2010). *Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative, and mixed methods* (3rd ed.). Sage Publications.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*. 13(3), 241-256. <https://doi.org/10.1177/026553229601300302>
- Nugroho. (2009). *UN dan Mutu Pendidikan [UN – the National Examination – and Education Quality]*. Suara Merdeka, 7 Desember 2009.
-

- Onaiba, A. M. (2013). *Investigating the Washback Effect of a Revised EFL Public Examination on Teachers' Instructional Practices, Materials and Curriculum* (Doctoral dissertation). University of Leicester. <https://hdl.handle.net/2381/28561>
- Orafi, S. M. S. & Borg, S. (2009). Intentions and realities in implementing communicative curriculum reform. *System*. 37, 234-253. <https://doi.org/10.1016/j.system.2008.11.004>
- Othman, I., Md Saleh, N., & Mohd Nooraini, N. A. (2013). The implementation of school-based assessment in primary school standard curriculum. *International Journal of Education and Research*. 1(7), 1-10.
- Pallant, J. (2010). *SPSS Survival Manual: A set by step guide to data analysis using SPSS* (4th ed.). McGraw-Hill.
- 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. <http://theses.gla.ac.uk/id/eprint/81300>
- Rea, L. M. & Parker, R. A. (2014). *Designing and Conducting Survey Research: A Comprehensive Guide* (4th ed.). Jossey-Bass.
- Saif, S. (2006). Aiming for positive washback: a case study of international teaching assistants, *Language Testing*. 23(1), 1-34. <https://doi.org/10.1191/0265532206lt322oa>
- Saukah, A. & Cahyono, A. E. (2015). National Exam in Indonesia and Its Implications to the Teaching and Learning of English. *Jurnal Penelitian dan Evaluasi Pendidikan [Journal of Research and Evaluation in Education]*. 19(2), 243-255. <https://doi.org/10.21831/pep.v19i2.5583>
- Shohamy, E. (2001) *The power of tests: A critical perspective on the uses of language tests*. Longman.
- 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>
- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers, *Educational Researcher*. 20(5), 8-11. <https://doi.org/10.3102/0013189X020005008>
- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1991). *The role of testing in elementary schools* (CSE Technical Report 321). University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Sutari, V. R. (2017). National Examination in Indonesia and Its Backwash Effects: Teachers' Perspective. *Atlantis Press*. 331-333. <https://doi.org/10.2991/conaplin-16.2017.76>
- Teddlie, C. & Tashakkori, A. (2009). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approach in the Social and Behavioral Sciences*. Sage.
- Wall, D. (2005). *The impact of high-stakes examinations on classroom teaching: a case study using insights from testing and innovation theory*. Cambridge University Press.
- Wall, D. & Alderson, J. C. (1993) Examining washback: the Sri Lankan impact study. *Language Testing*. 10, 41-69. <https://doi.org/10.1177/026553229301000103>
- Watanabe, Y. (2004). Methodology in Washback Studies. In Cheng, L. Watanabe, Y., and Curtis, A. (Eds.), *Washback in Language Testing: Research Context and Methods* (pp. 19-36). Lawrence Erlbaum Associates.
- Xu, Q. & Liu, J. (2018). *A Study on the Washback Effects of the Test for English Majors (TEM): Implications for Testing and Teaching Reforms*. Springer. <https://doi.org/10.1007/978-981-13-1963-1>
-

Ziegler, N. & Kang, L. (2016). Mixed methods design. In Moeller, A. J., Creswell, J. W., & Saville, N. (Eds.), *Second Language Assessment and Mixed Methods Research* (pp. 51-83). Cambridge University Press. <https://doi.org/10.4324/9781315747293-10>

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2024 Author(s). This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.