

## **High Stakes Testing Cancellation and Its Impact on EFL Teaching and Learning: Lessons from Indonesia**

**Ashadi**

Universitas Negeri Yogyakarta, Indonesia  
*Email: ashadi@uny.ac.id*

**Margana**

Universitas Negeri Yogyakarta, Indonesia  
*Email: margana@uny.ac.id*

**Siti Mukminatun**

Universitas Negeri Yogyakarta, Indonesia  
*Email: sitimukminatun@uny.ac.id*

**Amrih Bekti Utami**

Universitas Negeri Yogyakarta, Indonesia  
*Email: bektiutami@uny.ac.id*

*Received: 26 June 2022*

*Reviewed: 27 October 2022-28 December 2022*

*Accepted: 29 December 2022*

*Published: 30 December 2022*

### **Abstract**

The differential impacts of high stakes testing (HST) on curriculum, pedagogy, students learning, teacher professional development, and equity have been known in the literature, but its cancellation impact is not yet known. Situated in the post cancellation policy in Indonesian education system, this study seeks to explore the impacts of such policy change on schools, teacher pedagogical practices, and students' English as a Foreign Language (EFL) learning. A multiple case study design was employed in three different high schools involving 3 school leaders, 3 English teachers, and 9 students in semi-structured and group interviews. Their voices were thematically analysed and constantly compared across cases. The results demonstrate shifts on: (i) the individual student's and school's reorientation towards their own vision and mission, (ii) curriculum driven and more varied EFL instructional practices, and (iii) more self-regulated learning activities among certain teachers and students. Further and wider scale investigation is required to look deeper on how these shifts occur among larger groups of stakeholders.

Keywords: high stakes testing; cancellation policy; EFL learning; school stakeholders

### **Introduction**

The main educational reform movement of the past three decades is standards-based accountability that has been enacted through the needs for public accountability and the

implementation of high-stakes testing at schools (Carnoy, Elmore, & Siskin, 2003). Literature has recorded that the implementation of high-stakes testing for school quality has risen in numerous countries worldwide (Deville & Chalhoub-Deville, 2011). Afflerbach (2005, p. 153) notes that its popularity has come from the assumption that (a) its standardized nature equals to fair treatment on test takers; (b) the test is perceived to have ‘scientific aura’; and (c) it looks natural and familiar because of the annual implementation across levels of education. Geographically, standardized wide scale school tests have also been implemented in the UK, US, Australia, Chile, and Indonesia and, to a certain extent, brought about academic and political controversies.

Koretz (2017) criticises if such accountability testing is really an effective evaluation model to measure students’ achievement scores on large-scale, state-wide tests as a means of refining student performance, teacher pedagogic effectiveness, and school success. Specifically reviewing its impact on English language learning at schools, Acosta’s et al review (2020) indicates the effects of the high-stakes testing regime on learners’ opportunity to absorb language and on their access to academic content as it neglects the primacy and complexity of developing content-specific literacies and measuring learners’ knowledge and skills using standardized achievement tests. In Indonesia, Cahyono and Saukah (2015) evaluate the impact of such policy on school discriminatory practice in English language teaching while Ashadi and Rice (2016) underline its further impact on teachers’ access to professional learning. All the controversies and criticisms finally ended in 2020 when the ministry of Education and Culture halted the implementation of the policy and offered a new form of assessment which is not considered high-stake.

The cancellation came together with the severe impact of the COVID-19 pandemic that adds the complexity of school conditions. The outbreak effect on education has been felt at all levels, from nursery to higher education, as education institutions must face closures, implement social distancing, and make decisions about how to maintain instructions of all subjects under the new condition. In addition, it also has impacts on the implementation of wide scale assessment such as the post-cancellation of national examination at schools. While the practical conditions at schools are interesting to scrutinize, there is a dearth of literature and theory in the effect of such policy change.

Supovitz (2009, p. 215) noticed that the relevant theoretical foundations of high-stakes testing are believed to come from: motivational theory as testing can arouse a sense of extrinsic rewards or punishment, theory of alignment in educational systems that suggests curricula, pedagogy, and evaluation to reinforce each other, informational theory with the provision of test data for decision-making. These conceptual relations bring about high stakes testing as a tool for improving student achievement and public education and accept it as an effective instrument in education reform through standardized tests. However, when such tool is removed from the education system, the impact has never been investigated. This knowledge paucity opens a space for better understanding the issue of high-stake testing cancellation in Indonesia particularly and for wider audience worldwide.

## Literature review

The pros and cons of high-stakes testing

Advocates of high stakes testing believe that wide scale standardized examinations can provide advantages to individuals, schools, systems. Phelps (2019) argues that giving a higher stake to a test usually has a favourable impact on student achievement as it is likely to increase more motivation and efforts. A high-stakes test, according to Steger, Schroeders, & Gnambs’s

meta-analysis (2018), is also considered easy to administer as it can be administered by a third party. Further, supporters of such programs have argued that the examination provide advantages to schools, supporting the improvement of more effective classroom teaching, and better student achievement (Phelps, 2012). Such high stakes testing programs are also believed to allow comparability and transferability of results across different education systems and jurisdictions (Santiago, et al, 2011). Data from such programmes can also be used by schools and systems to assess how well they have taught different skills and competencies mandated in the curriculum, and to concentrate their staff development programmes on certain areas (Collier 2010). Therefore, proponents of HST tend to maintain its implementation despite protests and objections from different parties and stakeholders.

There are also studies showing the dangers of such program on different aspects of education such as curriculum, pedagogical practices, psychological aspects, equity, and even professional development. For examples, Amrein and Berliner's (2003) meta-analysis and other smaller scale studies of Klein et al. (2000) and Haney (2000) found that achievement gains following the implementation of high stakes testing programmes are mainly a result of teachers teaching to the test. Meanwhile, Ashadi and Rice (2016) suggest that such testing regimes could impact teacher segregation in terms of access to professional learning due to the narrowing curriculum that emphasizes certain tested subjects ahead of the others. Likewise, Jankowski (2020) notes the possibility of more inequity creation because of the more demanding tools and access for education. In terms of teacher professional development, Ananda and Ashadi (2021) discovered the diminishing teachers' role as materials developers as a result of teaching to the test caused by the implementation of high stakes testing.

The test results can frequently become a means by which schools are judged for the quality of education they deliver (West, 2010) for instance like NAPLAN in Australia. In a more detailed account, Lingard (2010, p.130) also warns the potential for the labelling of underachieving schools, which most likely are in poor communities and therefore, could neglect the strong relationship between socioeconomic status (SES) and student performance. On students, Peters and Oliver (2009) highlight a standardization tendency leading to neglection of differences in the needs, talents and achievements of varied students. In line with the previous the findings, Polesel, Rice and Dulfer (2014) report high degrees of student burden related to high-stakes tests in Australia. Learning-wise, Au (2008) criticizes the pervasiveness of high stakes testing to shape the students' educational experiences that limit the development of the range of today's knowledge and skills. He also argues that teachers tend to apply more teacher-centred instructional approaches in attempts to cover the contents and levels required by the tests. The sheer of international and domestic evidence has led to the government decision to end the standard based assessment policy and opened a door for new knowledge in what impacts such policy would have on schools, English teachers, and students.

### English as a foreign language in Indonesia

In Indonesia, English as one of the tested subjects in the previous high stakes testing regime held the significant status since secondary school despite its function as a foreign language. Lie (2007) criticizes the commitment to language competence as assigned by the curriculum and the race for high scores. The dilemma, according to Furaida et. al (2015) is more intense among low achieving schools with students from lower socio-economic status. Saukah and Cahyono (2015) confirm this school segregation and stress the possible adverse impact on teacher professional development.

Reflecting on the condition of English language teaching (ELT) in Indonesia, Larson (2014) suggests a critical pedagogy that teachers can aspire a more transformative approach in their professional practice. Policy-wise, Zein (2017) urges the government even further to be clear with the status of English as a foreign language in Indonesia as it could have impact on the practice of ELT in schools and future vision. Whereas the demand for English teaching and learning is increasing, Lauder (2008) voices the needs for a firmer policy so that instructional practices and outcomes can benefit the nation in general.

Wang & Hill (2011) claim that ELT teachers and professionals in Indonesia have been influenced by the instructional model developed and imposed by Anglo-speaking countries. The common practices of memorization and form-focused have been pervasively employed and although found to be ineffective. Curriculum change (Ashadi, 2015) adds the complexity of the issue as it entails subsequent teacher professional development programs that often problematic considering the size of the Indonesian education system. Whatever curriculum is implemented, teachers are always to blame when expected results do not come. In fact, research has shown that the swing of the wide scale standard-based assessment are stronger than the swing of the curriculum on teachers' instructional practice (Ashadi & Rice, 2016). Research as mentioned in the previous sections has also demonstrated empirical evidence on the danger and adversity of such assessment practice. It was not until the pandemic arrived that finally the assessment policy change happened.

In relation to the severity of the pandemic, the Indonesian ministry of education have cancelled the implementation of the National Examinations at schools since 2020. Isbell and Kremmel (2020) call for an agenda to ensure decision-making and security issues are well addressed if a high-stakes examination is to proceed during the pandemic. In the UK, the pandemic has also impacted on cancellation of education assessments at different levels, such as high school, pre-University national testing, graduating assessments, and professional examinations (UK Government, Department for Education 2020). Such cancellations of high-stakes testing may have influences on other aspects of education, particularly on the instructional practice held by teachers at schools.

So far, there has been no studies that investigate the impact of cancelling high stakes testing at schools in different education systems. It is not yet known what happens in the classroom pedagogical practice of English after the cancellation of the National Examination in Indonesia. There is still very little research in this field, especially related to the post-high stake testing era in the teaching of English in the classroom. It is also unknown to what extent the impact of the cancellation of high stakes testing policy on teachers' pedagogical practices and students' learning of English. To increase our understanding on the issue, we propose the following questions.

1. What do EFL teachers, school leaders, and students voice in response to the NE cancellation?
2. How do they make sense of the post HST experience?
3. How does the cancellation impact on EFL teachers' instructional practice?

## Method

Maxwell (2008) and Cohen et al. (2005) noticed the significance of research questions in line with research objectives, conceptual framework, design, and trustworthiness. With the formation of research questions designed to investigate the impact of NE on English instructional practice in high schools, the current study employed a qualitative multiple-case study design. Gerring (2004, p. 341), defined it as rigorous research of several units to comprehend a bigger unit

which is called Quintain by Stake (2006). This design allowed researchers and participants collaboration enabling the latter to tell their stories, encounters, and perspectives (Crabtree & Miller, 1999). Flyvbjerg (2006) added that such design could result in context-dependent knowledge and the studied cases could help the researchers' learning process of effective research practice and professional learning.

This study followed Crowe's et al (2011) steps to understand a case comprehensively: (1) defining the case; (2) selecting the case(s); (3) collecting and analysing the data; (4) interpreting data; and (5) reporting the findings. Based on the formulated research questions, this study focussed on how the NE cancellation has impacted on the English instructional practices at high schools in Indonesia. The selected cases were three public schools from different locations, status, and types to gain a thick description, rich information, and replication logic for theoretical generalization (Yin, 2009). The following step was deciding the school stakeholders as the participants who would provide in-depth information relevant to the case being studied. They must have the following criteria: (i) experiencing the policy change in their tenure, (ii) allowing different perspectives (administrative, pedagogic, and motivational) for a thick description, and (iii) representing school stake holding parties. From each school site, we selected purposively a school leader, an English teacher from grade XII who are assumed to know and experience a change due to the NE cancellation, and three student representatives to corroborate the previous accounts from different stakeholders.

Table 1. Research settings and participants

Case	Type	Location	Participants	Instruments
1	Public high school	Rural	1 Principal, 1 English teacher, 3 students of different gender, SES, and Achievement	Individual semi-structured interviews Group interview with students Instructional documents
2	Private high school	Urban	1 Principal, 1 English teacher, 3 students of different gender, SES, and Achievement	Individual semi-structured interviews Group interview with students Instructional documents
3	Vocational school	Sub urban	1 Principal, 1 English teacher, 3 students of different gender, SES, and Achievement	Individual semi-structured interviews Group interview with students Instructional documents

To gain information from these different school stakeholders, the research team constructed three sets of research questions previously validated by an expert in qualitative interviews. These were, then, the trialled in a pilot study involving three participants and the result was used to revise the instruments. We conducted a 60-minute individual semi-structured interview with every participating teacher and leader in the three sites. Participating students in each school were asked to give voice in a group interview to save time but contest potential differences. Relevant instructional and policy documents were asked and examined as a means of data corroboration and triangulation.

The interview records were then transcribed, read repeatedly to find relevant and interesting themes while other research members listened the recordings and noted in a table for

confirmation. Thematic analysis on each transcription was conducted paradigmatically based on the potentially emergent notions from the research questions. We followed Miles, Huberman, and Saldana (2014, p.10) who recommend moving *from one inductive inference to another by selectively collecting data, comparing, and contrasting this material in the quest for patterns or regularities, seeking out more data to support or qualify these emerging clusters, and then gradually drawing inferences from the links between other new data segments and the cumulative set of conceptualizations*. In this way, their three simultaneous stages of activity: data condensation, data display, and conclusion drawing/verification could facilitate the research process interactively.

In line with the multi-case nature, a cross-case analysis was conducted to compare what emerged and happened in one school case to the others. This study employed matrices and tables as suggested by Stake (2006) to do the cross-case comparison. Such comparative display facilitated the researchers to draw a bigger conclusion (Quintain) which is a result of iterative intra-case and cross-case comparisons.

## Findings

The findings of this study address three issues related to stakeholders' perceptions toward the abolition of the national examination and the influence on English teachers' pedagogical practice in school. The first is teachers' pedagogical experiences after the elimination of the National Examinations program. The next finding is concerned with their interpretation of the change in relation to their practices in the school. Finally, the findings sum up the areas of pedagogical practice that changed particularly in their last year schooling program post National Examination abolition policy.

### Teachers' pedagogical experiences post high stakes testing program

The abolition of National Examinations gave new perspectives to teachers in the way that they could choose what teaching approach suitable during the teaching-learning process. Even though the teachers came to the point at which the policy to alter the national exam might be blurred in some ways, they believed that under the guidance of all experienced stakeholders, the teachers could help students find the meaning of their learning. Moreover, high-stakes testing or National Examinations program in Indonesia stifle the reality of teachers' practices at the grassroots level. Therefore, the abolition of National Examinations gave the teachers a chance to adhere to curriculum driven. Here, every individual teacher's interpretation of the curriculum can be transformed into a more practical approach including students' active involvement during classroom activities and considering a range of materials in a broader spectrum - not in the narrow one as focused during a national-examination era. In brief, the experiences of different stakeholders can be summed up in the following emergent themes:

Table 2. Stakeholders' Voices

School leaders	Teachers	Students
- Reorientation of school priority	- Approach to students	- Focus to university entrance (HS) or vocational skills (VHS)
- English is still important	- Curriculum driven	- Less burden/ more relax
- Focus on curriculum	- Linguistics skills exploration	- Less motivated (perceived)
	- Limited time	
	- Role shift as motivators	

The students were used to be prepared to come to learning with preconceptions about the endpoint of National Examinations - benchmarking their academic achievement by simply taking the standardized test at a final stage of their study, focusing on how to achieve passing grades at few subjects, and drilling-based activities which seemed to be boring. The abolition of the National Examination has brought different feelings, responses, and practices as voiced in the following excerpts:

T1/Ivy/C1: *Perhaps what caused the difference was our efforts to maintain these students' motivation.*

Li/Ziggy/C2: *We need to redevelop their (students') learning motivation. The key is enjoyment. So, the hardest thing to do when no national exam is their lack of interest and engagement (in learning).*

However according to Izzy, some high-achieving students who are more self-regulated tend to be unaffected by the change. The teachers also believed that in the old days, their options to introduce linguistics skills to students were very limited in number, resulting in monotonous activities. However, in post national-examination era, teachers could provide students with some activities that varied since the teachers were welcome to explore linguistics skills other than reading certain text types, which was quite demanding back then. Nicky, an English teacher at a public high school noted, *with the abolition of the National Examination, the teaching strategy is capturing what teaching should be like, in accordance with national curriculum. In this case, for example, if the curriculum demands to improve students' theoretical skill, we will use cognitive activities. Likewise, if practical skills are requested, we will provide the students with practical activities, for instance, showing them some video tutorials.* The more room for varied explorations of teaching activities and methods have been voiced across cases but how teachers interpret such experiences through their roles and pedagogical practices can be different due to personal and contextual factors.

#### Making sense of EFL teachers' experiences after policy change

High-stakes tests have been used to change teaching and learning is a common practice in many countries in the world, especially in those with centralized educational systems (Shohamy et al., 1996). Such has undermined the role of teachers as agent of change although transferring knowledge and facilitating students during their academic endeavours in post national exam era would not automatically make them the agent of change in education. Furthermore, educational researchers studying the influence of high-stake testing on EFL teachers' professional development (Ashadi and Rice, 2016) have argued that this type of testing led to the constraints on professional growth to certain teachers and gave more privileges to the others.

Leading a vocational school, Ziggy emphasized: *So, with the abolition of the National Examination, schools are now more flexible in developing learning related to improving (students') vocational competence (Li/Ziggy/C2).* He added, *"Now we can refocus on the implementation of what we (Vocational Schools) have prepared together with the industrial and professional world"*. In the context of senior high school, Nicky clarified, *"We feel more capable of achieving the (school) target without being burdened, the children are also less stressed, and the teachers are not too stressed because of the school ranks"*. There is a sense of particular school identity in these voices which was previously uniformed by the implementation of high stakes testing.

The abolition of the National Examination provides a new perspective for teachers in choosing a suitable teaching approach during the teaching and learning process. Although teachers

have come to the point where the policy for changing national examinations may be vague in some respects, they believe that under the guidance of all experienced stakeholders, teachers can help students discover the meaning of their learning. Moreover, high-level examinations or the National Examination program in Indonesia hinder the reality of teacher practice at the grassroots level. Therefore, the abolition of the National Examination provides an opportunity for teachers to stick to the curriculum rather than the test. Here, each teacher's approach to align with the curriculum becomes more assertive and practical including students' active participation during class activities and considering a wide range of materials in it unlike the previous era that focused on the national exams.

Students are used to coming to study with preconceived notions about National Examination endpoints - comparing their academic performance with only taking standardized tests in the final stages of their studies, focusing on how to achieve passing grades in some subjects, and boring drilling-based activities. Teachers also believe that in the past, their options for introducing linguistic skills to students were very limited, resulting in monotonous activities. However, in the post-national exam era, teachers can provide students with several varied activities because teachers are welcome to explore linguistic skills apart from reading certain types of text, which at that time was quite demanding. Nicky, an English teacher at a public high school, noted that with the abolition of the National Examination, *the teaching strategy was to capture what teaching should be like, according to the national curriculum*. Succeeding that has been prescribed in the curriculum documents become teachers' main objectives. In this case, for example, if the curriculum demands to improve students' theoretical skills, we will use cognitive activities. Likewise, if practical skills are requested, we will provide students with practical activities, for example, showing them some video tutorials.

#### Pedagogical practices that changed in grade XII

In addressing the question of what pedagogical practices areas have changed in EFL classes especially that of 12 graders, the findings revealed that the answers were almost similar with a very slight difference. One of the facts that there was a shift in students' academic goals during their final year study should be of interest to address because the students' viewpoints have explicitly stood in the exact direction. The summary can be seen in the following table.

Table 3. Pedagogical practices

No	Features	During HST era	Post HST era
1	Approach	<i>Teaching to the test</i>	<i>Curriculum driven</i>
2	Teaching focus	<i>Receptive skills (reading)</i>	<i>More productive skills</i>
3	Time allocation	<i>More teaching hours (including extra test drills)</i>	<i>More time for language skills practice</i>
4	Assessment type	<i>Objective</i>	<i>Open-ended &amp; performance</i>

Even though staff leaders found the abolition of high-stake testing may not raise significant issues to teaching practices and students' academic success, one perspective from school leaders brought up the issue of students' motivation. Winny, the principal in the first school case, declared, *"Their (students') motivation to learn seems to decrease and cause difficulties for teachers to engage them in learning, moreover with the current pandemic condition. Some of them think that they will pass/graduate whether they learn or not (Li/Winny/C1)*. Such condition is shared across the schools and construed as the most difficult challenge faced by the EFL teachers when giving

instructions. Students, particularly, the low achievers or from lower SES tend to feel unmotivated as they realize that they may not continue their education to tertiary level.

In relation to EFL learning, teachers and students articulated the sense of more freedom in learning resulting in their expressions like: *video making, more collaboration and group works, focus on language skills development, and instructional exploration*. Nicky highlighted that *with the abolition, our teaching strategies tend to follow what the curriculum prescribes, if it is knowledge, we focus more on cognitive load and if it is skills, we just stick to the curriculum demand for example with video display, role play, or group work* (Ti/Nicky/C3). Meanwhile in the vocational school Ziggy added *they emphasize learning through projects and cases as (the teachers) have enough time when the school doesn't need to prepare for the National Examinations* (Li/Ziggy/C2). The shift to a more attractive and varied instructional approaches was confirmed by the participating students such as Abba and Talis after experiencing the last couple of years in their school.

Overall, there have been positive changes towards independence and alignment in curricular implementation, more productive language skill practice, varied forms of assessment, student self-regulation, and the accomplishment of school mission. Different stakeholders shared an almost parallel post-high stakes testing experience that was aggravated by the Covid-19 pandemic forcing EFL teachers to interact with students via online distance learning. The demotivation phenomena among students emerging in this study, hence, can possibly be due to the termination of high stakes testing policy or the online distance learning. A more valid and consistent claim can only be made when the learning condition is back to normal although the new normality is said to be like *putting the online genie back into its bottle* (Tesar, 2020).

## Discussion

To interpret and describe the significance of the findings, this section will discuss (i) how the school reorientate students' motivation post national examination cancellation, (ii) the freedom teachers feel and have in terms of instructional practice, and (iii) The specific pedagogical practice change. Each sub-section elaborates the respective issue in depth and breadth through comparing with relevant and current literature and theories to explain any new understanding or insights that emerged from the current research findings.

### Reorienting students' motivation

In line with Dowley and Rice (2022) this study found that a need for further examination, in broader samples and across age groups, of how high stakes tests have lifted the learning motivation of most students and how testing situation as well as the nature of the stakes for those being examined. Despite the claims from several teachers, students, and school leaders this study revealed that it is necessary to re-examine whether the high stakes have motivated students. In the case of high-school students, high-achieving participants seem to be able to reorientate their motivation towards university entrance test as their new stake.

Previous studies indicate that assigning substantial consequences to test performance stimulates students to study harder and learn extra (Angrist & Lavy, 2009; Braun et al., 2011). Research has also shown that the implementation of high stakes exams increased students' efforts, parental support, and school endeavours. Thus, high stakes tests are believed to lead high-performing students to increase their effort and achievement on tested subjects, but not for their low-performing counterparts. Simzar et al (2015), therefore, warns test practitioners to consider models of achievement motivation when proposing, implementing, and using the gained information from high-stakes assessments. Such consideration is important to make sure that

schools' and teachers' effects on student learning may be accurately measured, and additionally, all students should get the benefit from an assessment system intended to encourage learning and achievement.

Test anxiety increases in high stakes testing situations and can bring about more detrimental effects for students with high test anxiety like low achieving students in the current study. They just want to complete their K-12 education and rarely see themselves attending a further college education. These students also seem to have a performance avoidance goal orientation and are predominantly vulnerable to high test anxiety because of their limited access to higher education and their availability in the job market.

Further, Noddings (2003, p. 1) has claimed that "*Happiness and education are, properly, intimately connected. Happiness should be the aim of education, and a good education should contribute significantly to personal and collective happiness*". From such a view, a good education (including schooling) should, therefore, be concerned about students' academic learning and Subjective Well-Being (SWB) (Dello-Iacovo, 2009). Drawn from Diener's theory of SWB, Tian (2008) has theorized SWB in school as how students personally assess and emotionally experience their school lives including high stakes tests.

In relation to such schooling experience, Elliot and Church (1997) projected three types of goal orientations: mastery, performance-approach, and performance-avoidance. Students focusing on developing competence through an emphasis on learning, deep understanding, and improvement are categorized into a mastery goal orientation. While those with a performance-approach goal orientation are focused on the exhibition of knowledge and skills in comparison to others by trying to outperform their classmates. Still others, with a performance-avoidance goal orientation, are focused on avoiding looking incompetent and being left behind. The responses from the student participants in the current study demonstrate a similar pattern as some high-achieving students seem to remain motivated despite the elimination of the National Examination (NE). Whereas several others, according to the teachers, experience diminishing motivation and the rest seek to new challenge or stakes in the university entrance tests.

#### More independent instructional practice

Under testing regime, less time was spent on other curriculum areas and adjustment of pedagogical practice and curriculum content to mirror the tests. Polesel, Rice, and Dulfer (2014) found, in Australia, that the modification of teaching and curricular practices is in response to concerns regarding the use and reporting of test results and the potential impact on schools, teachers, and students. Meanwhile, in the current study, the concern was voiced more on how the use of the test results for the high school students' higher education accesses and opportunities. As a result, teacher instructional practice was driven more by the high stakes test items rather than by the curriculum content. With the abolition, teachers viewed an opportunity to head their teaching back to the curriculum and a space for creative teaching.

Further, according to Assaf (2008), testing pressures not only impact teachers' instructional practice and awareness to students' learning needs but can negotiate a teacher's professional identity and can influence teachers' accountability and ethical sense of what they should do for their students and who they need to be as teachers. The present study has shown how Issy could feel a shifting role more as a motivator in her sense of agency as English teachers during the forced online distance learning, post high stakes testing regime. Other participating teachers also report more instructional freedom in their efforts to succeed the curricular contents.

The emphasis of learning should, therefore, be on the individual learner language development, rather than on the assessment tool (test). Robinson and Dervin (2018) suggest dropping the number of high stakes testing and concentrating on traditionally related instruction to best keep schooling education. Therefore, what the participants do with refocussing on the curriculum is deemed justified as they want to give more attention on students' learning improvement. If stakeholders are required to seek accountability, they may look through portfolio-based assessments, school inspections, or sampling for national assessments as what the current Indonesian education system begins with its newly implemented minimum competency assessment (AKM).

This study was conducted during the Covid-19 pandemic which has caused schools, teachers, and students to remain work from home and do learning activities online. Such condition, together with the current educational policy change, may have impact on these stakeholders' perceptions and responses gathered as the main data in this research. The loosening regulations allowing for hybrid learning model and even face-to-face (traditional) learning may be perceived and made sense differently by them, and eventually their different responses. Additionally, considering the limited sample of high school level, this study might not capture the phenomenon occurring at lower schooling levels (elementary and junior) that might demonstrate a different tendency, for example with the persisting high stakes testing at local (district) levels. To gain more stable data and capture the phenomenon in more details, further research with wider coverage and larger participants can be conducted to confirm the findings of the current study.

#### Specific English language instruction change

Research suggest that high-stakes testing has mixed impacts on English language teaching. Menken (2006) found that high-stakes tests become de facto language policy in schools, and Solórzano (2008) showed high stakes tests as currently constructed are inappropriate for English language learners. However, Dikilitaş, Dollar and Mede (2018) found that English language teachers are often involved in the different stages of test preparation, and that this involvement has a positive impact on their testing knowledge and skills. Yet, such impact on students learning has been criticized due to the merely test-driven nature of students learning. Chen, Hao & Xiao (2020) noted that changes in the stakes of a language test in the Chinese university context had less impact on students' English learning motivation and approaches. While it was done in a different setting, the participants of the current study could demonstrate a more independent behaviour with their reorientation towards higher education entrance test as the new benchmark.

Regarded as non-skill-oriented assessment in English at different levels of education, HST neglects the potential of students in different language skills mastery (Sharjeel & Qazi, 2012). Such practice may bring about the notion amongst the students to rely on endeavours to pass English language examinations in written form only. The language-based competency and skills are neglected aspects under the high-stake testing regime as voiced by most of the participants. They currently enjoy teaching language based on the prescribed curriculum and developing more varied learners' language skills. However, the new orientation towards higher education entrance test has also become a new challenge, as it only focuses on specific reading skills.

Canagarajah (2006) argues that the risk to define proficiency based on a sole variety and difficulty to teach or measure it in several varieties simultaneously implies on the need to revise the dominant paradigms of assessment. The changing pedagogical priorities suggest a movement away from a dependence on discrete-item tests to instruments that are more sensitive to performance and pragmatics. In other words, language assessment would focus more on

transactional strategies, contextualized performance, communicative repertoire, and linguistic responsiveness. These are in line with what the participants of current study voiced through what they called as more objective and performance-based types of assessments.

## Conclusion

The abolition of the national examinations at schools has, to a certain extent, changed the orientation of the school management, the practice of EFL teaching conducted by teachers, and student's individual learning direction. Pedagogically, the changes involve closer alignment to the government prescribed curriculum, focus on more productive skills, and eventually lead to a more performance-based language assessment. The notion of self-regulated learning becomes a key in this new phenomenon as both teachers and students must decide what is best for their teaching/learning practice and self-development. Teachers realize they need to adjust to their new roles and approaches in the post-high stakes testing era with more alignment to the existing school curriculum, freedom for instructional exploration, and focus on productive language skills.

The change that has led to the perceived diminishing motivation among low-achieving students, at the same time, opened a space for EFL teachers to further follow up with more interesting and engaging language skills practice and more varied ways of assessing students' English language competencies in line with the curriculum requirements. With less pressure in the instructional practice after the policy abolition, it is important to support EFL teachers at every level of education, to innovate, explore, and stay updated on the latest pedagogical (theory, practice, and policy) development. However, further research is required to examine whether the higher education entrance test will emerge as a new stake that may drive English teachers' instructional practice and students' learning behaviour.

## Declaration of conflicting interest

The authors declare that there is no conflict of interest in this work.

## Funding acknowledgement

This study was fully funded by DIPA Universitas Negeri Yogyakarta B/1085/UN34.12/PT /2021 of Academic Year 2021.

## References

- Acosta, S., Garza, T., Hsu, H. Y., Goodson, P., Padrón, Y., Goltz, H. H., & Johnston, A. (2020). The accountability culture: A systematic review of high-stakes testing and English learners in the United States during No Child Left Behind. *Educational Psychology Review*, 32(2), 327-352.
- Afflerbach, P. (2005). National Reading Conference policy brief: High stakes testing and reading assessment. *Journal of Literacy Research*, 37(2), 151-162.
- Ananda, E., & Ashadi, A. (2021). High-stakes testing and English teachers' role as materials developers: Insights from vocational high schools. *Studies in English Language and Education*, 8(1), 115-130.
- Angrist, J., & Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American economic review*, 99(4), 1384-1414.
- Ashadi, A., & Rice, S. (2016). High stakes testing and teacher access to professional opportunities: lessons from Indonesia. *Journal of Education Policy*, 31(6), 727-741.

- Assaf, L. C. (2008). Professional identity of a reading teacher: responding to high-stakes testing pressures. *Teachers and teaching: theory and practice*, 14(3), 239-252.
- Au, W (2008a) Between education and the economy: high-stakes testing and the contradictory location of the new middle class, *Journal of Education Policy*, Vol. 23, No. 5, pp. 501-513.
- Au, W (2008b) Devising inequality: a Bernsteinian analysis of high-stakes testing and social reproduction in education', *British Journal of Sociology of Education*, Vol. 29, No. 6, pp. 639-651.
- Braun, H., Kirsch, I., & Yamamoto, K. (2011). An experimental study of the effects of monetary incentives on performance on the 12th-grade NAEP reading assessment. *Teachers' college record*, 113(11), 2309-2344.
- Canagarajah, S. (2006). Changing communicative needs, revised assessment objectives: Testing English as an international language. *Language Assessment Quarterly: An International Journal*, 3(3), 229-242.
- Carnoy, M., Elmore, R., & Siskin, L. (Eds.). (2003). *The new accountability: High schools and high-stakes testing*. New York: Routledge.
- Chen, Q., Hao, C., & Xiao, Y. (2020). When testing stakes are no longer high: impact on the Chinese College English learners and their learning. *Language Testing in Asia*, 10(1), 1-18.
- Cohen, L., Manion, L., & Morrison, K. (2005). *Research methods in education (5 ed.)*. London: Routledge.
- Crabtree, B. F., & Miller, W. L. (1999). *Doing qualitative research (2 ed.)*. Thousand Oaks, CA: Sage Publications.
- Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC medical research methodology*, 11(1), 1-9.
- Dello-Iacovo, B. (2009). Curriculum reform and 'quality education' in China: An overview. *International journal of educational development*, 29(3), 241-249.
- Deville, C., & Chalhoub-Deville, M. (2011). Accountability-assessment under No Child Left Behind: Agenda, practice, and future. *Language Testing*, 28(3), 307-321.
- Dikilitaş, K., Dollar, Y. K., & Mede, E. (2018). English language teachers' perceptions of involvement in high-stakes testing in Turkey. In *Teacher Involvement in High-Stakes Language Testing* (pp. 233-244). Springer, Cham.
- Dowley, M., & Rice, S. (2022). Comparing student motivations for and emotional responses to national standardised tests and internal school tests: The devil in the detail. *Australian Journal of Education*, 66(1), 92-110.
- Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of personality and social psychology*, 72(1), 218-232.
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219 - 245.
- Furaidah, F., Saukah, A., & Widiati, U. (2015). Washback of English national examination in the Indonesian context. *TEFLIN Journal*, 26(1), 36-58.
- Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review*, 98(2), 342-354.
- Isbell, D. R., & Kremmel, B. (2020). Test review: Current options in at-home language proficiency tests for making high-stakes decisions. *Language Testing*, 37(4), 600-619.
- Jankowski, N. A. (2020). *Assessment during a crisis: Responding to a global pandemic*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

- Koretz, D. (2017). *The testing charade: pretending to make schools better*. Chicago: University of Chicago Press.
- Larson, K. R. (2014). Critical pedagogy (ies) for ELT in Indonesia. *TEFLIN journal*, 25(1), 122-138.
- Lie, A. (2007). Education policy and EFL curriculum in Indonesia: Between the commitment to competence and the quest for higher test scores. *TEFLIN journal*, 18(1), 01-15.
- Lingard, B. (2010). Policy borrowing, policy learning: Testing times in Australian schooling. *Critical studies in education*, 51(2), 129-147.
- Maxwell, J. A. (2008). Designing a qualitative study. In L. Bickman & D. Rog (Eds.), *The handbook of applied social research methods* (2 ed.). Thousand Oaks CA: Sage Publications.
- Menken, K. (2006). Teaching to the test: How No Child Left Behind impacts language policy, curriculum, and instruction for English language learners. *Bilingual Research Journal*, 30(2), 521-546.
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook*. Los Angeles: Sage.
- Mukminin, A., Eddy, H., Makmur, M., Failasofah, F., Fajaryani, N., & Thabran, Y. (2013). The achievement ideology and top-down national standardized exam policy in Indonesia: Voices from local English teachers. *Turkish Online Journal of Qualitative Inquiry*, 4(4), 19-38.
- Noddings, N. (2003). *Happiness and education*. Cambridge: Cambridge University Press.
- Peters, S., & Oliver, L. A. (2009). Achieving quality and equity through inclusive education in an era of high-stakes testing. *Prospects*, 39(3), 265-279.
- Phelps, R. P. (2012). The effect of testing on student achievement, 1910–2010. *International Journal of Testing*, 12(1), 21-43.
- Phelps, R. P. (2019). Test Frequency, Stakes, and Feedback in Student Achievement: A Meta-Analysis. *Evaluation Review*. 43 (3–4): 111–151. doi:10.1177/0193841X19865628.
- Polesel, J., Rice, S., & Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: A teacher perspective from Australia. *Journal of education policy*, 29(5), 640-657.
- Robinson, B. K., & Dervin, A. (2019). Teach to the student, not the test. *Urban Education Research & Policy Annuals*, 6(2).
- 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*, 19(2), 243-255.
- Sharjeel, M. Y., & Qazi, W. (2012). Why does Testing English Language Skills Really Matter? Issues and Challenges in Pakistani Higher Education. *Journal of Educational Research*, (1027-9776), 15(1). 93-107.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language testing*, 13(3), 298-317.
- Simzar, R. M., Martinez, M., Rutherford, T., Domina, T., & Conley, A. M. (2015). Raising the stakes: How students' motivation for mathematics associates with high-and low-stakes test achievement. *Learning and individual differences*, 39, 49-63.
- Solórzano, R. W. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of educational research*, 78(2), 260-329.
- Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, 10(2–3), 211–227. <https://doi.org/10.1007/s10833-009-9105-2>.

- Steger, D., Schroeders, U., & Gnambs, T. (2018). A Meta-Analysis of Test Scores in Proctored and Unproctored Ability Assessments. *European Journal of Psychological Assessment*, 36: 1–11. doi:10.1027/1015-5759/a000494.
- Tesar, M. (2020). Towards a post-Covid-19 ‘new normality?’: Physical and social distancing, the move to online and higher education. *Policy Futures in Education*, 18(5), 556-559.
- Tian, L. (2008). Developing scale for school well-being in adolescents. *Psychological Development and Education*, 24(3), 100-106.
- UK Government, Department for Education. (2020). Coronavirus (COVID-19): cancellation of GCSEs, AS and A-Levels in 2020. [London (UK)]: Government of UK; [accessed 2020 June 30]. <https://www.gov.uk/government/publications/coronavirus-COVID-19-19-cancellation-of-gcse-as-and-a-levels-in-2020/coronavirus-COVID-19-19-cancellation-of-gcse-as-and-a-levels-in-2020>
- Wang, H. & Hill, C. (2011). A Paradigm Shift for English Language Teaching in Asia: From Imposition to Accommodation. *The Journal of ASIA TEFL*, 8 (4), 205-230.
- West, A. (2010). High stakes testing, accountability, incentives, and consequences in English schools. *Policy & Politics*, 38(1), 23-39.
- Yin, R. K. (2009). *Case study research: Design and methods* (5<sup>th</sup> Ed). Los Angeles: Sage.
- Zein, M. S. (2017). Elementary English education in Indonesia: Policy developments, current practices, and future prospects: How has Indonesia coped with the demand for teaching English in schools?. *English Today*, 33(1), 53-59.