



ORIGINAL ARTICLE

Assessing the Washback Effects of the 2025 National High School Graduation Exam: A Cross-Sectional Study in Vietnam

Chi Le Thi Nguyen^{1,+},
Quyen Tieu Nguyen Dao²

¹The University of Adelaide, Australia;

²Hung Vuong High School for the Gifted, Vietnam

⁺Corresponding author • Email: lechi.nguyengl@gmail.com

Article history

Received: 02 February, 2025

Accepted: 23 November, 2025

Published: 28 December, 2025

Keywords

Washback effect, assessment validity, Bloom's taxonomy, national high school graduation exam, curriculum reform

ABSTRACT

In the context of ongoing educational reforms in Vietnam and the growing emphasis on competency-based assessment, this study evaluates the Newly Proposed National High School Graduation Exam 2025 in English and its washback effects on students' learning. The evaluation investigates three key aspects: the alignment between the National High School Graduation Exam 2025 and the Ministry of Education and Training's 2018 General Education English Curriculum, the application of Bloom's taxonomy in test item design, and the exam's washback effects on teaching and learning practices. A case study was conducted at a high school in Vietnam with a sample of 150 students, employing a quantitative method. The findings reveal notable misapplications of Bloom's taxonomy in training materials, creating potential confusion for teachers in classifying and designing test items. Furthermore, the limited construct validity of the exam - particularly the exclusion of listening and speaking skills - has resulted in adverse washback effects, such as narrowing the choice of learning content and reducing student engagement. Meanwhile, the exam exerts a strong extrinsic motivational effect by driving students to achieve higher scores for university admission. The study contributes to the ongoing debate on curriculum-assessment alignment in Vietnam and highlights both the strengths and weaknesses of the National High School Graduation Exam 2025. It concludes with recommendations for policymakers and educators to enhance assessment validity, improve teacher training, and promote balanced washback effects that support communicative competence and long-term language learning.

1. INTRODUCTION

English teaching and learning in Vietnam have experienced notable changes over the years, influenced by various socioeconomic, political, and cultural factors (Vu, 2020). Between 1982 and 2002, English was a compulsory subject in high school level and an elective in lower secondary school levels, resulting in the development of the 7-year and 3-year programs, respectively. These programs primarily used grammar-based textbooks (Hoang, 2010), which emphasized grammar-translation methodology and behaviorist approaches (Vu, 2020). However, these programs did not yield learners with sufficient English proficiency, prompting calls for curriculum reform, leading to a significant initiative based on the Ministry of Education and Training (MOET)'s policy of "one curriculum, multiple textbooks."

As part of this reform, the MOET launched the 2018 General Education English Curriculum (GEEC), which serves as the national framework for English language teaching. Unlike previous practices where a single textbook dictated the curriculum nationwide, the GEEC offers flexibility by allowing various publishing houses to develop

their coursebooks within this framework. The GEEC aims for students to develop linguistic, sociolinguistic, discourse, and strategic competencies (Hoang, 2022). The goal for the 10-year English program is for learners to achieve Level 3 on the “Foreign Language Proficiency Framework of Vietnam” (MOET, 2014), equating to B1 on the CEFR scale. To meet this goal, the curriculum emphasizes communicative methods and learner-centered pedagogy, transforming the role of teachers into instructors, mentors, and co-participants in the learning process. Consequently, students are encouraged to engage actively and collaboratively in their own education (Hoang, 2022).

This shift in pedagogical focus is also reflected in assessment practices. The GEEC moves away from traditional assessments predominantly centered on grammar and reading, requiring evaluations that align with communicative learning outcomes. This includes listening, speaking, reading, and writing competencies, along with pronunciation, vocabulary, and grammar, and promotes ongoing assessment through both formative and summative methods.

The year 2025 is significant as it marks the completion of the 3-year high school GEEC program by the first cohort of students. In light of this, the MOET has put forward the Newly Proposed National High School Graduation Exam 2025 (NPNHSGE 2025) for the English subject. This new exam format is intended to significantly differ from prior assessments by shifting the emphasis towards meaningful language use in context rather than isolated grammatical tasks. However, early observations indicate that the test may still focus predominantly on grammar and reading, raising concerns regarding its consistency with the GEEC objectives and potential effects on teaching and learning practices.

Despite the crucial role of the national high school graduation exam in the Vietnamese education system, research on its washback effects is notably limited. This study aims to fill that gap by exploring the alignment of the NPNHSGE 2025 with the GEEC, the application of Bloom’s taxonomy in its test items, and the exam’s effects on students’ motivation, engagement, and learning strategies. The evaluation will take place at Hung Vuong High School for the Gifted, where the authors are located.

By contextualizing the NPNHSGE 2025 within the ongoing reforms in English education in Vietnam, this article aims to enrich scholarly discourse on curriculum-assessment alignment and shed light on the implications of large-scale testing for teaching and learning practices in the Vietnamese educational landscape.

2. LITERATURE REVIEW

2.1. Bloom’s Taxonomy and Reading Comprehension

The original taxonomy developed by Bloom and his colleagues provides systematic definitions for six major categories in the cognitive domain: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. Except for the application category, each of these categories is further divided into subcategories. Bloom’s taxonomy is regarded as one of the most influential frameworks in curriculum studies, evidenced by its wide application, which has generated approximately 7,350 results in the Google Scholar database. This practical influence underscores its relevance in the field of education.

It is important to note that Bloom did not design the taxonomy solely as an assessment tool. His primary intention was to create a common language for articulating learning goals, providing a foundation for translating general objectives into specific outcomes, and aligning objectives with teaching, learning, and assessment practices.

This conceptualization is exemplified in Table 1, which presents a representative list of verbs associated with each cognitive category (Munzenmaier & Rubin, 2013).

Table 1. Bloom’s taxonomy verb list

	Skill	Definition	Verbs
Level 1	Knowledge	Recall information	Identify, describe, name, label, recognize, reproduce, follow
Level 2	Comprehension	Understand the meaning, paraphrase a concept	Summarize, convert, defend, paraphrase, interpret, give examples

Level 3	Application	Use the information or concept in a new situation	Build, make, construct, model, predict, prepare
Level 4	Analysis	Break information or concepts into parts to understand it more fully	Compare/contrast, break down, distinguish, select, separate
Level 5	Synthesis	Put ideas together to form something new	Categorize, generalize, reconstruct
Level 6	Evaluation	Make judgments about value	Appraise, critique, judge, justify, argue, support

Source: Munzenmaier & Rubin (2013)

While the taxonomy offers clear distinctions among cognitive levels, ambiguities persist, particularly in the domain of reading comprehension. Many educators classify “reading to answer questions” as a comprehension skill. However, as Beatty (1975) pointed out, this practice often aligns more closely with the knowledge level, as it typically involves identifying or recalling explicit details rather than engaging in deeper interpretive processes.

Such ambiguity is critically relevant when evaluating the NPNHSGE 2025. Although its designers claim that test items fall within three categories - Knowledge, Comprehension, and Application - a closer examination may suggest alternative classifications. This raises important questions about the validity of the exam’s design and opens up opportunities for improving assessment practices, inspiring hope for the future of education.

2.2. Validity and Washback Effects

In the realm of educational testing, validity is an essential concept that evaluates how suitable a test is for its intended purpose. One relatively new and exciting facet of validity is alignment. This idea focuses on how well national educational assessments align with the standard curricula across various countries, highlighting the importance of providing relevant and practical evaluation tools (Geisinger, 2013).

When there is a lack of validity - particularly a disconnect between the curriculum’s learning outcomes and the test’s objectives - the consequences can be significant. This phenomenon is known as the washback effect, which can significantly influence both teaching practices and student learning experiences (Xu & Liu, 2018). In fact, the washback effect can be so powerful that it often shapes instructional methods and educational approaches.

A wealth of influential research has delved into the effects of washback, uncovering both its positive and negative implications for educators and learners alike. Alderson and Wall laid the groundwork in 1993 by proposing fourteen insightful hypotheses that help illuminate the intricate nature of washback effects (Xu & Liu, 2018). They emphasized that because washback is multifaceted, understanding it requires consideration of the test’s characteristics alongside the specific educational contexts in which they are applied. Hughes (1993) added to this conversation by suggesting a framework that distinguishes among participants, operations, and outcomes in teaching and learning, enriching our comprehension of how washback operates (Xu & Liu, 2018). Building on these ideas, Bailey introduced a foundational model of washback in 1996 that elegantly merges the theories of Alderson and Wall with Hughes’ trichotomy (Xu & Liu, 2018).

Since then, numerous empirical studies have built upon this theoretical groundwork, often focusing on large-scale high-stakes exams, as these tend to have the most profound impact on language education. For example, Nguyen (2025) explored the washback effects of the IELTS exam, revealing that the test’s influence notably shaped both learning motivation and the experience of limited learning opportunities. In a similar vein, Doan and Piamsai (2025) examined the VSTEP.3-5 test and found that it significantly affected the learning strategies employed by undergraduate students.

However, it’s important to acknowledge a gap in research, particularly in Vietnam, regarding the washback effects of the NPNHSGE, one of the pivotal assessments in a student’s educational journey. This presents an exciting opportunity for further exploration and understanding in this essential area of educational assessment.

3. MATERIALS AND METHODS

3.1. Sampling

This study focused on a sample of 150 Year-12 students from five different classes at Hung Vuong High School for the Gifted. The relatively large sample size was intended to provide a comprehensive perspective from students with diverse academic abilities and various orientations towards higher education.

As Hung Vuong High School for the Gifted is located in an urban area, the participants' attitudes towards the national examination may be similar to those of students from other cities. To reduce potential bias from students majoring in English, two English-specialized classes were excluded from the sample. However, since Hung Vuong High School for the Gifted is a selective institution that admits students through competitive entrance examinations, the findings of this study should not be generalized to the entire student population in Vietnam.

3.2. Data Collection and Analysis

To address the research objectives content analysis was employed. The relevant materials included official dispatches from the GEEC issued by the MOET, the exam format for the NPNHSGE 2025, and teacher training documents. Additionally, the researchers accessed internal training resources to enrich the dataset.

The documents were systematically analyzed to identify key patterns in learning outcomes and the cognitive levels of test items. These findings were then compared with the General Education English Curriculum learning outcomes and teacher training materials to examine potential mismatches between the NPNHSGE 2025 and the General Education English Curriculum, as well as any misapplications of Bloom's taxonomy in classifying test items.

Additionally, an online questionnaire in Vietnamese was administered to the 150 students. Prior to the primary survey, a pilot study was conducted with 30 students to evaluate the clarity of the item wording and ensure that the questionnaire was comprehensible. The survey had no time limit, and the students were instructed to read each statement carefully and select the option that best reflected their views. After incorporating feedback from the pilot phase, the final questionnaire was distributed, resulting in 150 valid responses (36.7% male and 63.3% female).

4. RESULTS AND DISCUSSIONS

4.1. Result

This evaluation was conducted to gather empirical findings that could contribute to enhancing the proposed test's construct and, in turn, improve the quality of English teaching and learning at the high school level in Vietnam. The results focus on two critical aspects: the validity of the test construct and the application of Bloom's taxonomy in the design of test items.

4.1.1. Lack of Validity and Misapplication of Bloom's Taxonomy

The first significant finding highlights the lack of validity in the construct of the NPNHSGE 2025. One of the exam's primary functions is to serve as a basis for high school graduation, which necessitates an assessment of whether students have reached Level 3 of the Vietnamese Foreign Language Proficiency Framework, equivalent to B1 on the CEFR scale. Achieving this level requires competence in all four skills: listening, reading, writing, and speaking. However, the exam design focuses solely on reading comprehension and grammatical knowledge. As a result, it is not possible to make valid inferences about students' proficiency in listening, speaking, and writing based solely on scores derived from reading and grammar tasks (see Appendix 1). This limited scope undermines the construct validity of the exam and raises concerns about whether the test outcomes accurately reflect the learning objectives of the GEEC.

The second significant finding pertains to the misapplication of Bloom's taxonomy in classifying cognitive levels. According to the official training document, test items were purportedly distributed across three categories: remembering, understanding, and applying, with designated proportional weights. However, the independent content analysis conducted for this study presents a different reality. Out of the 40 test items examined, 14 items fell into the remembering category. These items primarily require students to identify or recognize appropriate English phrases or expressions. Although some of the phrases may seem rare or unfamiliar to students - leading teachers or test designers to mistakenly classify them as higher-order thinking tasks - a closer inspection reveals that they operate only at the lowest cognitive level.

Moreover, 26 items were identified as belonging to the understanding category. These tasks ask students to extract information from reading passages, select correct answers, or paraphrase ideas. While such tasks may imply a deeper

level of processing, they primarily assess whether students comprehend the given material rather than their ability to apply knowledge in new contexts. Consequently, despite the difficulty of some reading passages or the use of B1-C1 vocabulary, the majority of the items still align with understanding rather than application. This distinction is critical because linguistic complexity or challenging vocabulary does not elevate a test item to a higher-order cognitive process, as defined by Bloom's taxonomy (see Appendix 2).

Together, these findings indicate that the NPNHSGE 2025, in its current form, fails to align with the intended outcomes of the GEEC regarding both construct validity and cognitive rigor. The exam design risks narrowing the focus of teaching and learning towards lower-order tasks, which could undermine the broader communicative and competency-based goals of the curriculum.

4.1.2. The Washback Effects on Learning Content, Engagement, and Motivation

The perception of students about the importance of the NPNHSGE 2025

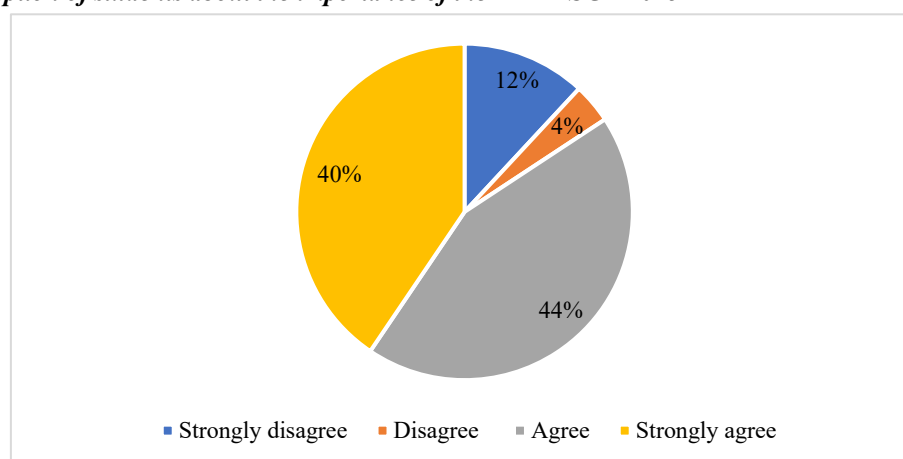


Figure 1. Students' perceptions of the importance of the NPNHSGE 2025 in shaping university choice and study objectives

Students' perceptions of the NPNHSGE 2025 highlight the significant washback effects that this exam has on their learning practices and academic goals. As illustrated in Figure 1, a total of 70.7% of the respondents (44% agreeing and 26.7% strongly agreeing) indicated that the outcome of this exam would significantly influence their choice of university. Only a small percentage expressed disagreement (3.3%) or strong disagreement (12%). These findings emphasize the critical role of the exam in shaping students' educational paths, as their performance on the test is viewed as closely related to future academic opportunities and career prospects (Bao & Cho, 2022).

Additionally, the survey results show that 64.6% of the students believe the exam provides them with clear learning objectives. They perceived the structure and requirements of the NPNHSGE 2025 as a roadmap for identifying which areas of language knowledge and skills to prioritize in order to gain admission to their desired universities. This perception suggests that the test has a strong motivational effect, aligning students' study strategies with the demands of the assessment.

However, while the exam appears to enhance students' focus and engagement, its effectiveness is highly influenced by how well the test content reflects the intended learning outcomes of the General Education English Curriculum. If the exam emphasizes only specific skills, students may dedicate disproportionate amounts of time and effort to those areas, potentially neglecting other competencies that are equally crucial for comprehensive language development.

The Washback Effects of the NPNHSGE 2025 on the Choice of Learning Content

Students' responses reveal significant washback effects of the NPNHSGE 2025 on their learning priorities and teachers' instructional focus. Table 2 summarizes the survey results.

Q3: Are the revision periods for the NPNHSGE in English necessary?

Q4: Should teachers focus mainly on topics that are directly related to the NPNHSGE 2025?

Q5: Do students want their teachers to go beyond exam-related content and cover wider topics in English?

Q6: Do teachers tend to emphasize test practice more than the four skills in the coursebook?

Q8: Do the sample tests provided by MOET help students in their preparation for the NPNHSGE 2025?

Q9: Do students find the mock tests provided by teachers similar to the NPNHSGE 2025?

Table 2. Students' perceptions of the washback effects of the NPNHSGE 2025 on learning content

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Majority Response
Q3. Revision periods are necessary for test preparation.	4.0%	7.3%	29.3%	47.3%	12.0%	59.3% agree/strongly agree
Q4. Teachers should focus on exam-related topics.	3.3%	6.0%	11.3%	47.3%	32.0%	79.3% agree/strongly agree
Q5. Teachers should cover topics beyond the exam.	12.7%	28.0%	20.0%	28.0%	11.3%	39.3% agree/strongly agree
Q6. Teachers emphasize test practice over the four skills.	4.7%	6.0%	33.3%	40.0%	16.0%	56.0% agree/strongly agree
Q8. MOET sample tests are helpful for preparation.	4.0%	7.3%	13.3%	31.3%	44.0%	75.3% agree/strongly agree
Q9. Teachers' mock tests are similar to the NPNHSGE 2025.	1.3%	9.3%	30.7%	30.7%	28.0%	58.7% agree/strongly agree

The results indicate that nearly 60% of the students either strongly agreed or agreed that revision sessions during school are essential for exam preparation (Q3). Additionally, 79.3% expected their teachers to focus on exam-related topics (Q4), while fewer than 40% believed that teachers should extend instruction to areas beyond the exam content (Q5). This shows that the exam significantly influences students' expectations, promoting a focus on test-specific material rather than broader communicative competencies.

Moreover, over 56% of the respondents reported that their teachers emphasized test practice more than the four language skills outlined in the coursebook (Q6). This highlights a shift in instructional priorities, where practice tests and exam drills are considered more valuable than skill-based activities like speaking or writing, which are not directly assessed in the exam.

The responses further reveal that 75.3% of the students found the MOET's official sample tests helpful for their exam preparation (Q8). Similarly, 58.7% felt that the mock tests created by their teachers resembled the actual NPNHSGE 2025 (Q9). These findings suggest that the exam not only directs students' learning behaviors but also influences teachers' instructional design, prompting them to align classroom activities closely with the exam format.

While this alignment may enhance students' immediate test performance, it also narrows the scope of learning. Skills such as listening, speaking, and writing - essential for the communicative goals of the General English Exam for Communication - receive less attention, limiting students' overall language development.

The Washback Effects of the NPNHSGE 2025 on Students' Engagement

Students' engagement with English learning under the influence of the NPNHSGE 2025 is summarized in Table 3.

Q10: Are the question items in the NPNHSGE 2025 related to topics that students find engaging?

Q11: Is the NPNHSGE 2025 designed to increase students' engagement with English learning?

Q12: Does practicing with mock tests for the NPNHSGE 2025 reduce students' engagement with English?

Table 3. Students' perceptions of the washback effects of the NPNHSGE 2025 on engagement

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Majority Response
Q10. Items in the test are related to engaging topics.	6.0%	10.7%	43.3%	25.3%	14.7%	43.3% neutral
Q11. The test increases engagement in English.	7.3%	48.7%	14.7%	25.3%	4.0%	56.0% disagree/strongly disagree
Q12. Practicing with mock tests reduces engagement.	10.0%	26.7%	26.7%	26.0%	10.7%	Responses evenly distributed

The results indicate that only 43.3% of the students felt that the test items were related to topics they found engaging (Q10). A similar percentage expressed neutrality on this matter, suggesting that for many students, the exam content did not resonate strongly with their personal interests. Additionally, a combined 56% disagreed or strongly disagreed with the statement that the exam was designed to enhance their engagement in English learning (Q11). This indicates that the exam format is not seen as motivating.

When it comes to the impact of practicing with mock tests, the student responses were mixed. Approximately 36.7% disagreed or strongly disagreed that practicing with mock tests reduced their engagement, while 36.7% agreed or strongly agreed (Q12). This balanced distribution suggests that mock test practice may have both positive and negative effects, depending on individual learners' experiences.

Overall, these findings highlight that while the NPNHSGE 2025 is regarded as academically significant, it does not consistently foster student engagement in English learning. Instead, the results point to a neutral or even negative washback effect on motivation, raising concerns about whether the test encourages sustained interest in language learning.

The Washback Effects of the NPNHSGE 2025 on Students' Motivation

Students' responses regarding their motivation to prepare for the NPNHSGE 2025 are presented in Table 4.

Q13: *Is students' motivation to prepare for the NPNHSGE 2025 derived from their interest in learning English?*

Q14: *Is students' motivation to prepare for the NPNHSGE 2025 derived from the pressure of obtaining high scores to gain admission to their target universities?*

Table 4. Students' perceptions of the washback effects of the NPNHSGE 2025 on motivation

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Majority Response
Q13. Motivation comes from interest in English.	8.0%	14.0%	40.0%	31.3%	6.7%	39.3% agree/strongly agree
Q14. Motivation comes from pressure to achieve high scores for university admission.	6.0%	21.3%	25.3%	46.7%	25.3%	72.0% agree/strongly agree

The results indicate a significant difference between intrinsic and extrinsic sources of motivation among students. Only 39.3% reported that their motivation to prepare for the exam came from a genuine interest in English, while a larger proportion (40%) expressed a neutral stance on the matter. In contrast, a substantial 72% acknowledged that their motivation was primarily driven by the pressure to achieve high scores for admission to their desired universities.

This suggests that while the exam exerts a strong motivational influence, it is largely extrinsically oriented. Students seem to be more affected by the high-stakes nature of the test and its impact on university opportunities than by a true enthusiasm for learning English.

4.2. Discussion

The findings of this study highlight two key issues: the validity of the NPNHSGE 2025 and its impact on teaching and learning practices. From a validity standpoint, the exam construct does not align with the objectives of the GEEC. By excluding listening and speaking skills, the exam undermines the communicative focus of the curriculum and raises questions about its effectiveness in representing students' actual language proficiency. Previous research (Ha & Tran, 2021) has shown that when essential skills are left out of high-stakes assessments, teachers are less likely to incorporate them into their classroom practices. This results in a situation where years of English instruction do not translate into real communicative competence, hindering national objectives for English language education.

Another critical issue is the cognitive demand placed on students. Although Bloom's taxonomy was intended to aid in classifying test items, this study found that its application has been inconsistent. Training materials often confuse linguistic difficulty with higher-order thinking, complicating the assessment process for teachers and leading to tasks that are unnatural and questionable from a pedagogical perspective. This misapplication not only confuses educators but also risks fostering assessment practices that are disconnected from genuine language use. It highlights a broader challenge in aligning theoretical frameworks with practical test design, where clarity of interpretation is essential to avoid oversimplifying Bloom's taxonomy into a mere labelling exercise.

The washback effects identified in this study are varied. On the positive side, the exam provides students with clear objectives and a sense of direction in their learning. It also motivates them to work hard in pursuit of high scores, which are crucial for university admission. This illustrates the exam's role as a significant extrinsic motivator within Vietnam's education system, consistent with Baker et al.'s (2013) observations regarding the motivational force of high-stakes assessments.

Conversely, the exam also narrows both teaching and learning. Teachers often prioritize exam-related topics, test preparation, and lower-order language tasks, sidelining broader communicative skills. This tendency supports Alderson and Wall's (1993) foundational claim that tests dictate what and how teachers instruct. As a result, the curriculum becomes more exam-driven than learner-centered, limiting students' opportunities to engage meaningfully with the language. Furthermore, the study suggests that the exam does not consistently promote engagement or intrinsic interest in English, echoing Feinman's (2008) assertion that high-stakes assessments can induce anxiety and lead to disengagement when students view them as removed from authentic communicative purposes.

In summary, these findings highlight the dual nature of washback. While the NPNHSGE 2025 offers motivation and clarity of purpose, it simultaneously reinforces a narrow, exam-focused learning culture that may undermine the broader objectives of language education. The challenge for policymakers and educators is to design assessments that balance accountability and fairness while ensuring alignment with curricular goals and fostering genuine communicative competence.

4.3. Recommendations

The findings of this study indicate that reforming the construct and implementation of the NPNHSGE 2025 is crucial for enhancing its validity and washback effects. Although Vietnam faces limitations in infrastructure and human resources, incremental yet meaningful adjustments can be introduced. One of the most urgent steps is to incorporate listening and writing skills into the exam construct. Excluding these skills not only diminishes the exam's validity but also discourages teachers and learners from fully engaging in developing communicative competence.

To address challenges related to the reliability of assessing productive skills such as writing, developing detailed analytic rubrics is essential. Systematic training and calibration sessions for examiners should accompany this. Such measures would reduce subjectivity in scoring and promote fairer assessment practices.

In terms of test design, the study suggests exercising caution when applying Bloom's taxonomy mechanically to multiple-choice questions in reading and listening. While Bloom's taxonomy is a valuable pedagogical tool, the classification of item types in large-scale examinations should consider difficulty levels and cognitive processing demands, rather than equating linguistic complexity with higher-order thinking. For an exam of national significance, a rigorous item piloting and validation process should be institutionalized prior to official administration. Piloting can help identify problematic items, calibrate difficulty levels, and ensure alignment with intended learning outcomes.

Beyond technical reforms, the study emphasizes the importance of shaping the perceptions of both teachers and students. To avoid narrowing the curriculum to test-oriented content, policy interventions should highlight that the ultimate goal of ten years of English instruction under the GEEC is to build communicative competence for academic and professional contexts. This requires awareness-raising initiatives, professional development programs, and curricular guidance that consistently reinforce the value of English as a means of communication rather than merely as an examination subject. Without changes in perceptions, even technically improved exams risk reinforcing test-driven teaching and learning.

5. CONCLUSION

This study offers valuable insights into the validity and washback effects of the NPNHSGE 2025, framed within the 2018 GEEC. The analysis identifies key challenges related to construct validity, the application of Bloom's taxonomy, and the exam's impact on students' motivation, engagement, and learning choices. These findings contribute to the ongoing discourse on curriculum and assessment alignment in Vietnam, providing evidence that can assist policymakers and test designers in refining the structure and implementation of high-stakes language assessments.

However, several limitations should be noted. The study was conducted with a relatively small sample from Hung Vuong High School for the Gifted, which restricts the representativeness of the findings. The participants do not reflect the diverse socio-economic and educational contexts across Vietnam, making it difficult to generalize the results nationwide. Additionally, the reliance on self-reported survey data introduces an element of subjectivity that may influence the reliability of the responses. Lastly, since the study was carried out within three months of the exam's release, it captures only immediate perceptions and short-term washback effects, without considering potential long-term impacts on teaching and learning practices.

Despite these limitations, the study represents a timely contribution by providing an evidence-based evaluation of the first implementation phase of the NPNHSGE 2025. The findings illuminate both the strengths and weaknesses of the exam while offering suggestions for reform, particularly in enhancing validity, clarifying assessment frameworks, and mitigating negative washback. Therefore, this study serves as a reference point for future research and policy development aimed at ensuring that national assessments support fairness and align with the broader educational objectives of English language teaching in Vietnam.

Conflict of Interest: No potential conflict of interest relevant to this article was reported.

REFERENCES

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129. <https://doi.org/10.1093/applin/14.2.115>
- Baker, B. D., Oluwole, J. O., & Green, P. C., III. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the Race-to-the-Top era. *Education Policy Analysis Archives*, 21, 5. <https://doi.org/10.14507/epaa.v21n5.2013>
- Bao, N. V., & Cho, Y. (2022). How the high-stakes and college entrance exam affects students' perception: Implication on management policy in higher education. *East Asian Journal of Business Economics (EAJBE)*, 10(2), 83-94. <https://doi.org/10.20498/eajbe.2022.10.1.83>
- Beatty, R. (1975). Reading comprehension skills and bloom's taxonomy. *Reading World*, 15(2), 101-108. <https://doi.org/10.1080/19388077509557305>
- Doan, P. A. N., & Piamsai, C. (2025). Washback of Vietnamese Standardized Test of English Proficiency (VSTEP.3-5) on Undergraduate students' learning strategies. *LEARN Journal: Language Education and Acquisition Research Network*, 18(1), 23-53. <https://doi.org/10.70730/vezk2760>
- Feinman, J. (2008). *High Stakes, but Low Validity? A Case Study of Standardized Tests and Admissions into New York City Specialized High Schools*. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit.

- Geisinger, K. F. (2013). *APA handbook of testing and assessment in psychology vol 1, test theory and testing and assessment in industrial and organizational psychology* (vol 1, pp. 61-84). American Psychological Association.
- Ha, X., Tran, N., & Tran, N. (2021). Teachers' beliefs and practices regarding assessment in English as a Foreign language Classrooms in Vietnam. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2021.5063>
- Hoang, V. V. (2010). The current situation and issues of the teaching of English in Vietnam. *立命館言語文化研究*, 22(1), 7-18.
- Hoang, V. V. (2022). Interpreting MOET's 2018 general education English curriculum. *VNU Journal of Foreign Studies*, 38(5), 1-22.
- Hughes, A. (1993). *Backwash and TOEFL 2000*. Unpublished Manuscript, Reading, England: University of Reading.
- Ministry of Education and Training (MOET, 2014). *Circular No. 01/2014/TT-BGDDT dated January 24, 2014 on promulgating the six-level Foreign Language Proficiency Framework for Vietnam*.
- Ministry of Education and Training (MOET, 2018). *Circular No. 32/2018/TT-BGDDT dated December 26, 2018 on promulgating general education program*
- Munzenmaier, C., & Rubin, N. (2013). *Bloom's taxonomy: What's old is new again*. The Elearning Guild, Santa Rosa.
- Nguyen, T. M. (2025). IELTS Washback as a High-Stakes Test on Student Learning: A hierarchical modelling study at a Vietnamese university. *International Journal of TESOL & Education*, 5(1), 1-24. <https://doi.org/10.54855/ijte.25511>
- Vu, T. T. (2020). *English language curriculum reform at the national level: A case of intentions and realities in Vietnam*. Published Ph.D. thesis, The University of Adelaide.
- Xu, Q., & Liu, J. (2018). A study on the washback effects of the test for English majors (TEM). In *Springer eBooks*. Springer Nature. <https://doi.org/10.1007/978-981-13-1963-1>

APPENDIX

Appendix 1. Comparison between the learning outcomes of GEEC and NPNHSGE

	General Education English Curriculum	National High School Graduation Exam
Upper Secondary Level Objectives	<p>Upon completing the upper secondary English curriculum, students are expected to:</p> <ul style="list-style-type: none"> • Use English as a means of communication through the four skills of listening, speaking, reading and writing so as to meet basic and practical communicative needs on familiar topics related to school, leisure, entertainment, employment and daily life. • Further develop core linguistic knowledge, including phonology, vocabulary and grammar; and, through English, gain broader understanding of English-speaking countries and other cultures worldwide; recognise and respect cultural diversity; and begin to express Vietnamese cultural values appropriately in English. • Use English to support learning in other school subjects. 	<p>Students are expected to demonstrate that they:</p> <ul style="list-style-type: none"> • Understand the main ideas of clear, standard texts or spoken input on familiar topics related to school, work and leisure. • Manage most communication situations that may arise in settings where English is used. • Produce simple connected texts on familiar topics or matters of personal interest. • Describe experiences, events, dreams, hopes and ambitions, and briefly justify opinions or plans.

	<ul style="list-style-type: none"> • Use English to pursue further study or employment after upper secondary graduation. • Apply effective learning strategies and time-management skills; make informed use of information technology for learning and self-study; develop autonomy and self-assessment; take responsibility for learning outcomes; and cultivate lifelong learning habits.
Expected Competency Level	<p>Students completing the GEEC are expected to achieve Level 3 of the Vietnamese Six-Level Foreign Language Proficiency Framework (equivalent to CEFR B1), meaning that learners can understand the main points of clear, standard communication on familiar matters; handle most situations likely to arise when travelling in areas where English is spoken; write simple connected texts on familiar topics; and describe experiences and ambitions while briefly explaining opinions and plans.</p> <p>Cultural awareness and personal qualities: Through English learning, students develop broader cultural understanding, respect for diversity and the capacity to represent Vietnamese cultural values in English; alongside personal qualities such as integrity, responsibility, empathy and national pride.</p>
	<p>Skills Assessed</p> <p>Reading: assessed through comprehension tasks focusing on familiar and contemporary topics.</p> <p>Listening: not assessed in the examination.</p> <p>Writing: not assessed in the examination.</p> <p>Speaking: not assessed in the examination.</p>

Appendix 2. Analysis of the NPNHSGE 2025 based on Bloom's Taxonomy

No.	Verb	Remembering	Understanding	Application
1	Identify	x		
2	Identify	x		
3	Identify	x		
4	Recognize	x		
5	Recognize	x		
6	Recognize	x		
7	Identify	x		
8	Recognize	x		
9	Recognize	x		
10	Recognize	x		
11	Identify	x		
12	Recognize	x		
13	Interpret		x	

14	Interpret	x
15	Interpret	x
16	Interpret	x
17	Interpret	x
18	Interpret	x
19	Interpret	x
20	Interpret	x
21	Interpret	x
22	Interpret	x
23	Identify	x
24	Paraphrase	x
25	Paraphrase	x
26	Paraphrase	x
27	Paraphrase	x
28	Interpret	x
29	Classify	x
30	Classify	x
31	Interpret	x
32	Paraphrase	x
33	Paraphrase	x
34	Identify	x
35	Summarize	x
36	Paraphrase	x
37	Interpret	x
38	Paraphrase	x
39	Interpret	x
40	Summarize	x

Appendix 3. Students' Questionnaire

Please indicate the extent to which you agree with the following statements regarding the NPNHSGE 2025 in English.

For each item, select one response:

A. Strongly Agree B. Agree C. Neutral D. Disagree E. Strongly Disagree

1. The result of the NPNHSGE in English has a significant influence on my choice of university.

2. The NPNHSGE in English helps me set clear learning goals for myself.

3. In-class revision sessions for the NPNHSGE in English are necessary.
4. I want my teachers to focus mainly on topics that are directly related to the NPNHSGE in English.
5. I want my teachers to extend instruction beyond the topics that are directly related to the NPNHSGE in English.
6. Teachers should concentrate more on practising test papers rather than developing all four language skills in the coursebook.
7. If the official English examination follows the sample tests issued by the Ministry of Education and Training, it will be very difficult.
8. The sample tests issued by the Ministry of Education and Training are helpful for my preparation for the NPNHSGE in English.
9. The practice tests used by my teachers are similar to the sample tests issued by the Ministry of Education and Training.
10. The topics included in the NPNHSGE in English are relevant to my personal interests.
11. The NPNHSGE in English is designed in a way that increases my engagement in learning English.
12. Practising mock tests for the NPNHSGE in English decreases my interest in learning English.
13. My motivation to prepare for the NPNHSGE in English comes from my genuine interest in learning English.
14. My motivation to prepare for the NPNHSGE in English comes from the pressure to achieve high scores for university admission.