

Test Perceptions and Preparation Strategies in the Chinese EFL

Context: A Washback Study

Huirong Geng¹, Harmi Izzuan Bin Baharum^{2*}

Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Kuala Lumpur Campus,
Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia

*Corresponding author

DOI: <https://dx.doi.org/10.47772/IJRISS.2025.910000772>

Received: 01 November 2025; Accepted: 07 November 2025; Published: 24 November 2025

ABSTRACT

This study investigates the relationships between Chinese EFL learners' perceptions of the National Matriculation English Test (NMET) and their preparation strategies. Drawing on survey data from 620 senior high school students, analyses examined how perceptions of test importance and test difficulty influence five categories of strategies: cognitive, metacognitive, memory, compensation, and social-affective. Results show that perceived importance is positively correlated with all strategies, particularly metacognitive regulation, suggesting that high stakes motivate structured and reflective preparation. In contrast, perceived difficulty correlated negatively with most strategies, especially cognitive and compensation, indicating that overwhelming challenge suppresses adaptive engagement.

Keywords: Washback, Test perceptions, Preparation strategies, NMET, Chinese EFL learners

INTRODUCTION

High-stakes examinations wield significant influence over language teaching and learning globally. In the realm of language education, this influence is referred to as washback (Alderson & Wall, 1993; Bailey, 1996; Messick, 1996). Existing research has clearly shown that such tests have a profound impact on curricula, teaching materials, and classroom practices (Cheng, 2005; Green, 2007a, b). However, the effects of these tests are not homogenous; they are moderated by the perceptions and actions of various stakeholders, with learners playing a particularly crucial role (Qi, 2005; Xie & Andrews, 2013).

Learners' perceptions of a test are fundamental in understanding their preparation strategies. Drawing from the expectancy-value and control-value theories (Eccles & Wigfield, 2002; Pekrun, 2006), students who attach high importance to a test are more likely to engage in activities such as study planning, self-monitoring, and self-regulation. Empirical studies have indicated that positive perceptions of test design tend to encourage the use of strategies that are relevant to the test construct (Xie & Andrews, 2013), while negative perceptions often lead to a narrowing of learning to test-wise tactics (Qi, 2007; Cheng, 2008). Despite a substantial body of literature on the washback effects of these tests, relatively few studies have systematically explored how learners' perceptions shape their preparation strategies.

This article aims to fill this research gap by investigating the relationship between test perceptions and preparation strategies among Chinese English as a Foreign Language (EFL) learners. The expected findings are intended to contribute to the development of washback theory, offer valuable insights into learner strategy use,

and provide practical implications for language teaching, educational policy-making, and test design.

LITERATURE REVIEW

NMET

Currently, the Chinese education system includes preschool, primary, secondary and tertiary education levels. Primary school education, which lasts six years, and junior secondary education, which lasts three years, are compulsory. After the completion of this nine-year compulsory education program, approximately 50% of junior high students enter senior high school for an additional three years of study. At the end of the senior high school education, the students are required to take the National College Entrance Examination (NCEE, known as Gaokao in China). NCEE has become one of the most important systems of choosing talents in the world. And scores gained in examination are used to evaluate students' academic performance. The National Matriculation English Test (NMET) is the English test of the National College Entrance Examination (NCEE).

In China, NMET has a great influence on students' future and destiny. For those Chinese students who can gain a high score by good performing will get chance to enroll in national key universities, while others may go to common ones or even lose opportunities of being admitted by universities. Due to its critical role in determining university placement, the NMET exerts considerable influence on English language instruction and learning at the secondary level. In many secondary schools, the final year of English instruction is largely devoted to NMET preparation. This often involves intensive test-focused activities, such as repeated practice with mock examinations, targeted exercises aimed at enhancing specific skills, and detailed error analysis.

Washback

Washback, also referred to as backwash, describes the influence that tests exert on teaching and learning. In language education, this concept has been widely examined because high-stakes assessments often shape what is taught, how it is taught, and how learners approach their studies. Alderson and Wall's (1993) seminal work formally conceptualized washback, proposing 15 hypotheses to explain the ways examinations affect educational practices. Building on this, Hughes (1993) introduced a trichotomy model, distinguishing between participants (teachers, learners, administrators), processes (teaching and learning activities), and products (learning outcomes). This model highlighted the channels through which tests exert their effects. Bailey (1996) extended the discussion by offering a basic model of washback that incorporated the direction and scope of test influence. She emphasized the interplay between the test, the participants, and the broader educational context, while distinguishing between positive and negative washback. Green's (2007a) model investigates washback from three dimension including washback direction, washback variability and washback intensity. The washback intensity, Green's third dimension, is measured in terms of the test importance and the test difficult. In order to explain the relationship between test difficulty, perceived test importance, and observed washback, Green developed a conceptual model. According to Green, there is strong washback when the test is viewed as significant and difficult. And there is a nonlinear link between perceived test difficulty and washback. Washback is produced at the midway point of the test difficulty continuum rather than at either extreme. Participants typically quit tests that they feel are too challenging, and they wouldn't bother studying if they thought the test was too easy. Intense washback is unlikely to happen unless the test is viewed as difficult yet achievable. This study will investigate what the senior high school students' perceptions of the NMET are, in terms of test difficulty and test importance, according to Green's Model of Washback (2007a).

Learning Strategies

According to Oxford (1990), language learning strategies are classified into direct strategies (memory strategies, cognitive strategies and compensation strategies) and indirect strategies (metacognitive strategies, affective

strategies, and social strategies). “Memory strategies, such as grouping or using imagery, have a highly specific function: helping students store and retrieve new information. Cognitive strategies, such as summarizing or reasoning deductively, enable learners to understand and produce new language by many different means. Compensation strategies, like guessing or using synonyms, allow learners to use the language despite their often large gaps in knowledge”. (p. 37) “Metacognitive strategies allow learners to control their own cognition, that is, to coordinate the learning process by using functions such as centering, arranging, planning, and evaluating. Affective strategies help to regulate emotions, motivations, and attitudes. Social strategies help students learn through interaction with others”. (p. 135)

In the present investigation, however, affective and social strategies are consolidated into a unified category labelled “social-affective strategies”, as participants’ engagement with these strategies was both limited and frequently overlapping.

Empirical Study

This part will discuss the empirical studies on the washback effects of high-stakes language assessments, exploring the perception of participants and teaching or learning process.

Qi’s (2005) work is crucial for showing the conflict between intended and actual washback. She found that despite test designers’ intentions to promote positive learning, learners’ and teachers’ perception of the test as a hurdle led to a learning process focused solely on test-taking tricks and practice exams, undermining deeper language acquisition. Students’ primary aim was simply to achieve a high score for university admission, reinforcing the emphasis on exam performance over language development.

Hawkey (2006) extended the inquiry to the International English Language Testing System (IELTS), exploring how perceptions of test fairness and validity mediate washback. Hawkey’s findings revealed a conditional pattern: when IELTS was viewed as fair and valid, washback tended to be constructive; when it was seen as overly mechanical or high-stakes, practices leaned toward drilling and test-oriented teaching. Importantly, Hawkey noted that learners consistently prioritized short-term performance gains over long-term proficiency development, a trend attributed to the existential stakes of IELTS results for their future opportunities.

Green’s work on IELTS further elaborated on these dynamics. In an in-depth investigation of IELTS writing test washback among pre-higher education learners (Green, 2007a), he found that two interrelated factors drove exam-focused preparation: test design features and learners’ perceived stakes. These factors led learners to prioritize skills directly tied to test performance, such as adhering to preset essay templates or addressing “key points” in writing prompts, over developing transferable writing abilities. In a related comparative study, Green (2007b) contrasted the learning engagement of learners in IELTS preparation courses with those in university pre-session English courses, which emphasize communicative competence over test performance. Findings confirmed that perceived test stakes strongly mediated engagement: learners in IELTS courses, who viewed the test as critical to accessing higher education, narrowed their focus to exam-specific strategies. By contrast, pre-session learners, with less immediate pressure to achieve a specific test score, engaged more broadly with communicative activities. This divergence underscored how high-stakes test perceptions can redefine “learning success” as test performance, rather than language proficiency.

Cheng (2008) expanded on these themes by examining high-stakes language tests globally. Classroom observations across contexts revealed that teaching practices became narrowly focused on test-taking strategies, past paper practice, and rote memorization.

Xie and Andrews (2013) refined these insights through a focused investigation of China’s College English Test (CET), a nationwide exit-level examination for university students. Combining classroom observations with post-test learner reflections, they revealed how specific design features differentially mediate preparation

behavior. Multiple-choice sections were interpreted by examinees as rewarding “test-wiseness” (e.g., distractor elimination) rather than deep lexical or syntactic knowledge; consequently, learners favored strategy drills over extensive reading or systematic vocabulary expansion. In the writing component, the highly specified rubric and predictable prompt genres encouraged the rote memorization of prefabricated introductions and argumentative schemata, effectively displacing opportunities for original composition and rhetorical flexibility. Xie and Andrews thus demonstrated that within a single high-stakes assessment, divergent task formats can elicit distinct, surface-level preparation strategies, each reinforcing a disjuncture between test performance and the broader communicative aims nominally espoused by the examination.

Razavipour et al. (2020) investigated the relationship between test-takers’ perceptions and their preparation behaviors within a washback framework. The study explored how learners’ views of test content and test uses influenced the intensity and nature of their preparation. They found that positive perceptions of test content were associated with more intensive preparation, suggesting that learners invest greater effort when they believe the test meaningfully reflects their skills. Interestingly, while perceptions of test uses were correlated with the value learners attached to test taking, these perceptions did not significantly predict actual preparation intensity. This highlights a nuanced pattern: learners may recognize the stakes and value of a test but do not always translate this recognition into concrete preparation behaviors.

Liu and Yu (2021) explored the relationships between motivational orientations and learning practices among 258 Chinese undergraduate English majors preparing for a high-stakes English test. Findings demonstrated a clear divergence in the effects of motivational types. Extrinsic test-driven goals were found to significantly and positively predict rote-oriented practices such as test drilling, repetitive memorization, and exam-focused rehearsal. By contrast, intrinsic goals, reflecting personal interest and value in language learning, were associated with broader and more balanced learning activities, suggesting more meaningful engagement with English beyond test preparation.

Notwithstanding these valuable contributions, several gaps in the literature remain. First, while existing research has extensively documented teacher practices and classroom-level washback, comparatively fewer studies have undertaken an in-depth exploration of learners’ own subjective perceptions. Second, although general strategic tendencies have been identified, there is limited empirical evidence detailing the specific categories of strategies that learners prioritize or neglect in direct relation to their perceptions. Finally, despite the global focus on tests like IELTS and TOEFL, the Chinese EFL context, dominated by the monumental influence of the NMET on millions of annual test-takers, warrants more systematic investigation.

To address these gaps, the present study seeks to explore the interplay between learner perceptions and strategic preparation within the context of the Chinese NMET. This study aims to examine how perceptions of the test’s importance and difficulty correlate with the reported use of distinct strategic categories. By investigating these relationships, this research endeavors to contribute to a more nuanced and learner-centered understanding of washback mechanisms, offering significant implications for pedagogical practice in high-stakes assessment environments.

METHOD

This study employed a quantitative design to examine the extent to which EFL learners’ perceptions of the National Matriculation English Test (NMET), specifically regarding its importance and difficulty, influence their adoption of specific preparation strategies.

Participants were final-year senior high school students (aged 17–19) from five senior high schools in Guangxi province, China. This cohort was selected because, having completed approximately three semesters of NMET

preparation and being imminent test-takers, they possessed direct and recent experience with NMET-focused learning activities, ensuring the relevance and recency of their responses. A total of 689 questionnaires were distributed, with 620 valid responses retained for analysis. The sample included 336 male participants (54.2%) and 284 female participants (45.8%), reflecting a relatively balanced gender distribution. Geographically, the inclusion of students from five high schools across Guangxi enhanced the sample's diversity, offering a broader perspective on NMET perceptions and preparation strategies among EFL learners in the region.

Data were collected using two questionnaires: a student perception questionnaire, designed to assess learners' views on NMET importance and difficulty, and a test preparation questionnaire, aimed at exploring their specific preparation practices. The perception questionnaire comprised two main sections. Section A gathered demographic information, while Section B included 12 items measuring perceptions of NMET importance and difficulty. Items related to importance used a 5-point Likert scale (1 = "does not apply to me at all" to 5 = "applies to me entirely"), and those assessing difficulty used a 5-point scale (1 = "very easy" to 5 = "very difficult"). The test preparation questionnaire contained 50 items, also using a 5-point scale (1 = "never" to 5 = "always"), to quantify the frequency of specific preparation behaviors.

Reliability analysis was conducted to evaluate the instruments' internal consistency. Following Perry Hinton et al. (2004), who note that a Cronbach's alpha coefficient between 0.70 and 0.90 indicates high reliability, all dimensions of both questionnaires yielded alpha values above 0.7, confirming their high reliability.

RESULTS AND DISSCUSSION

Demographic Data

Participants report that they have learned English for a duration ranging from 9 to 17 years, with an average of 9.63 years. The participants spent between 0 and 4 hours daily on learning English, with an average of 0.87 hours per day. For the last-year students in senior high school, they have to prepare for the other 5 subjects like Chinese and math. The participants have prepared for the NMET for an average of 2.79 semesters, with a range from 1 to 4 semesters. This suggests that most participants have engaged in a structured preparation process over 3 semesters. On average, schools administer 2.92 tests per month for students, with a range from 1 to 4 tests, indicating that participants experience frequent testing. Participants report taking an average of 2.96 tests per month independently, with a range from 0 to 20 tests. This wide range suggests that while some participants engage in extensive self-testing, others may not test themselves at all.

Descriptive Patterns

Table 1 presents the descriptive statistics and Pearson correlation coefficients examining the relationships between learners' perceptions of the importance and difficulty of the NMET, with their reported use of five categories of preparation strategies (N = 620).

Table 1 Correlation Between Importance, Difficulty and Preparation Strategies

Strategy	Mean	Std. Deviation	Importance		Difficulty	
			Pearson Correlation (r)	Significance (p-value)	Pearson Correlation (r)	Significance (p-value)
Cognitive	3.063	0.661	0.128**	0.001	-0.220**	< 0.001
Compensation	3.368	0.747	0.187**	< 0.001	-0.237**	< 0.001
Memory	3.620	0.699	0.173**	< 0.001	-0.091*	0.023
Metacognitive	3.014	0.502	0.257**	< 0.001	-0.052	0.194
Social-Affective	3.113	0.762	0.188**	< 0.001	-0.103*	0.01

Notes: ** indicates significance at the 0.01 level (2-tailed). * indicates significance at the 0.05 level (2-tailed).

Descriptive statistics revealed that among the five strategy categories, memory strategies were employed most frequently ($M = 3.62$, $SD = 0.70$), followed by compensation strategies ($M = 3.37$, $SD = 0.75$). In contrast, metacognitive strategies yielded the lowest mean score ($M = 3.01$, $SD = 0.50$), slightly below the scale midpoint, indicating that learners engaged in fewer planning, monitoring, and evaluation activities compared to practice- and rehearsal-oriented strategies.

The relatively high frequency of memory strategy use aligns with the enduring emphasis on rote learning and memorization in Chinese exam-focused educational contexts, as noted in prior research on NMET preparation. In Confucian-heritage educational settings, memorization is not only a traditional approach but also a socially legitimized strategy, particularly in high-stakes exam contexts. This suggests that memory strategies may be broadly employed regardless of learners' test perceptions.

Perceived Importance and Preparation Strategies

Correlation analysis indicated that learners' perceptions of NMET importance were positively associated with all five categories of preparation strategies, with statistically significant correlations observed for cognitive ($r = .128$, $p < .01$), compensation ($r = .187$, $p < .001$), memory ($r = .173$, $p < .001$), metacognitive ($r = .257$, $p < .001$), and social-affective strategies ($r = .188$, $p < .001$). Studies on high-stakes exams, such as the TOEFL, IELTS, and China's NMET, have consistently found that students who perceive a test as highly important tend to invest more time and effort in structured test preparation (Xie & Andrews, 2013). Green (2007b) found that IELTS candidates who considered the test critical for academic progression engaged more in metacognitive regulation, while Qi (2004) observed that Chinese learners preparing for the NMET adopted goal-setting, error-tracking, and repeated rehearsal routines when the stakes were perceived as high. The finding of this study aligns with this pattern, as students who perceive the NMET as significant are more likely to engage in cognitive, metacognitive, memory, compensation, and social-affective strategies to improve their performance.

Metacognitive strategies exhibited the strongest relationship with perceived importance ($r = .257$, $p < .001$), suggesting that learners who attach greater significance to the NMET are particularly inclined to engage in regulatory behaviors such as planning, monitoring, and evaluating their preparation efforts. The strong relationship between test importance and metacognitive strategy use is widely recognized in washback research. High-stakes language tests, such as the NMET, IELTS, and TOEFL, promote greater metacognitive engagement because students recognize the necessity of planning, monitoring, and evaluating their learning progress (Cheng, 2005; Green, 2007b). Xie and Andrews (2013) found that Chinese EFL learners who viewed the NMET as critically important adopted more structured and goal-oriented study behaviors, such as setting learning objectives, tracking progress, and regulating test-taking approaches. The strongest association reflects a tendency for students to channel their motivation into self-regulatory processes. Learners who regard the NMET as highly important appear not only to practice more but also to prepare more deliberately through careful planning of study schedules, self-monitoring of weaknesses, and systematic evaluation of progress.

Perceived Difficulty and Preparation Strategies

Correlation analysis revealed significant negative relationships between learners' perceptions of NMET difficulty and several categories of preparation strategies. Specifically, perceived difficulty was negatively correlated with cognitive ($r = -.220$, $p < .001$), compensation ($r = -.237$, $p < .001$), memory ($r = -.091$, $p < .05$), and social-affective strategies ($r = -.103$, $p < .05$). These findings indicate that as learners perceive the NMET to be more difficult, their reported use of these strategies decreases. Several studies also indicated that when students perceive a test as too difficult, they may feel overwhelmed and reduce their engagement in strategy use. Xie and Andrews (2013) found that Chinese EFL learners preparing for the NMET who perceived the test as highly difficult were less likely to engage in active learning strategies, as they felt that their efforts would not

lead to success. Cheng (2005) also noted that high perceived test difficulty can discourage learners from using a variety of strategies, as they may feel that no amount of preparation can significantly improve their performance.

In contrast, the relationship between perceived difficulty and metacognitive strategies was not statistically significant ($r = -.052$, $p > .05$), suggesting that learners' regulatory and monitoring processes remained relatively unaffected by their perceptions of test difficulty. Metacognitive strategies are not significantly influenced by perceived difficulty, meaning students may continue to plan and monitor their learning regardless of difficulty. Research on washback effects has shown that metacognitive strategies tend to remain stable across different testing conditions, as learners who use these strategies do so as part of their broader academic habits rather than as a reaction to test difficulty. Spratt (2005) found that while test difficulty influenced cognitive and compensation strategies, metacognitive strategies remained unchanged, further supporting this finding.

CONCLUSION

The findings of this study reveal that NMET preparation is shaped by learners' perceptions in multifaceted ways. The relatively high frequency of memory strategy use aligns with the enduring emphasis on rote learning and memorization in Chinese exam-focused educational contexts. Learners who attach greater significance to the NMET are particularly inclined to engage in regulatory behaviors such as planning, monitoring, and evaluating their preparation efforts. And Metacognitive strategies exhibited the strongest relationship with perceived importance. Students who view the NMET as more difficult may be discouraged from actively employing the most strategies, but the relationship between perceived difficulty and metacognitive strategies was not statistically significant

These findings yield several salient implications. For learners, the findings indicate the need to move beyond reliance on memorization by incorporating more reflective practices such as systematic planning, self-monitoring, and ongoing evaluation of progress. For teachers, the results suggest the importance of explicitly demonstrating the value of strategies that extend beyond rote learning. For schools and test-preparation programs, the study highlights the value of integrating explicit training in metacognitive strategies, such as time management and reflective evaluation, into their curricula. For policymakers, the evidence underscores that perceptions of high test difficulty may discourage students from engaging with effective strategies. Consequently, it is essential to provide accessible preparation resources and clear guidance that build learners' confidence in the utility of strategic approaches to exam preparation.

Although this study has yielded valuable insights into the relationships between learners' perceptions of the NMET and their preparation strategies, several avenues for future research warrant exploration.

Qualitative exploration of learner perceptions

This study relied on quantitative survey data and correlation analysis to identify overarching patterns. However, qualitative methods, such as semi-structured interviews and learner diaries, could offer deeper insights into how learners interpret test importance and difficulty, and how these perceptions shape their day-to-day preparation practices. Such approaches would enrich understanding of the washback mechanism at the individual level.

Longitudinal investigations

Future research should examine how learners' perceptions evolve across different stages of exam preparation, from the early years of senior secondary school to the final months preceding the NMET. A longitudinal design would reveal whether perceptions of importance and difficulty fluctuate over time and how these changes correspond with shifts in strategy use, preparation intensity, and eventual performance outcomes.

Strategy quality versus frequency

While the present study measured the frequency of strategy use, future work should consider the quality and effectiveness of strategy application.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the participating senior high schools for facilitating the questionnaire process. Their cooperation was essential to the successful completion of this study.

REFERENCES

1. Aldson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics*, 14(2), 115-129.
2. Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language testing*, 13(3), 257-279.
3. Cheng, L. (2005). *Changing language teaching through language testing : a washback study*. Cambridge University Press. Publisher description <http://www.loc.gov/catdir/enhancements/fy0664/2006277397-d.html>
4. Cheng, L. (2008). The key to success: English language testing in China. *Language testing*, 25(1), 15-37.
5. Dong, M. (2020). Structural relationship between learners' perceptions of a test, learning practices, and learning outcomes: A study on the washback mechanism of a high-stakes test. *Studies in Educational Evaluation*, 64, 100824.
6. Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology*, 53(1), 109-132.
7. Green, A. (2007a). *IELTS washback in context: Preparation for academic writing in higher education (Vol. 25)*. Cambridge University Press.
8. Green, A. (2007b). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education*, 14(1), 75-97.
9. Hawkey, R. (2006). *Impact theory and practice (Vol. 24)*. Cambridge University Press.
10. Hughes, A. (1993). *Backwash and TOEFL 2000*. Unpublished manuscript, University of Reading.
11. Liu, X., & Yu, J. (2021). Relationships between learning motivations and practices as influenced by a high-stakes language test: The mechanism of washback on learning. *Studies in Educational Evaluation*, 68, 100967.
12. Messick, S. (1996). Validity and washback in language testing. *Language testing*, 13(3), 241-256.
13. Oxford, R. L. (1990). Language learning styles and strategies. *Georgetown University Round Table on Languages and Linguistics (GURT) 1990: Linguistics, Language Teaching and Language Acquisition: The Interdependence of Theory, Practice and Research*, 438.
14. Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, 18(4), 315-341.
15. Perry Hinton, D., Hinton, P. R., McMurray, I., & Brownlow, C. (2004). *SPSS explained*. Routledge.
16. Qi, L. (2004). *The intended washback effect of the National Matriculation English Test in China: Intentions and reality*. Foreign Language Teaching and Research Press.
17. Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language testing*, 22(2), 142-173.
18. Qi, L. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education*, 14(1), 51-74.

-
19. Razavipour, K., Mansoori, M., & Shooshtari, Z. G. (2020). Test takers' perspectives on an English language test in Iranian higher education: A washback study. *Issues in Educational Research*, 30(3), 1058-1083.
 20. Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language teaching research*, 9(1), 5-29.
 21. Xie, Q., & Andrews, S. (2013). Do test design and uses influence test preparation? Testing a model of washback with Structural Equation Modeling. *Language testing*, 30(1), 49-70.