

RESEARCH

Open Access



How sociocultural factors mediate washback effect: a study of the Hong Kong Diploma of Secondary Education Examination English Language Paper 4

Nozomu Wong Katagiri^{1*} 

*Correspondence:
k634364@kansai-u.ac.jp

¹ Kansai University, 3-3-35
Yamate-Cho, Suita-Shi, Osaka,
Japan

Abstract

The study explores the unaddressed washback on the English-speaking test of the Hong Kong Diploma of Secondary Education Examination (HKDSE-English Language Paper 4). The HKDSE is a university entrance exam in Hong Kong which includes a group discussion speaking test known as English Language Paper 4. To identify ongoing washback effects, including sociocultural perspectives, the study analyses ongoing student washback of test-specific and non-specific preparation and how the mediating factors affect these preparations. One-hundred and ninety-nine secondary 4th to 6th grade students provided the questionnaire data in two rounds. Results from exploratory factor analysis indicate six major types of test preparation, which are categorised as two test specific, three non-specific, and one integrated type of preparation. Mediating factors comprise five categories of learner characteristics and stakeholders in and out of school. The cluster analyses identified four students' groups intertwined with the involvement of washback and the mediating factors. A strong relationship between extrinsic factors and explicit learning and intrinsic factors and implicit learning was found. The study concludes that students' washback is affected by mediating factors aside from the test; learning method choice by students (e.g. learning from entertainment content) has a strong relationship with specific mediating factors (e.g. interest in English).

Keywords: Washback, Hong Kong diploma secondary education, English Language Paper 4, Extrinsic, Intrinsic, Mediating, Cluster analysis, Sociocultural, Learning

Introduction

The examination of the sociocultural contexts around test-takers and how they mediate students' learning have been widely adopted in recent washback studies (e.g. Cheng et al., 2011; Tsang & Isaacs, 2022; Zhan & Andrews, 2014). From the first washback model purposed by Alderson and Wall (1993), decades of ongoing studies (e.g. Bailey, 1996; Booth, 2012; Green, 2006; Saville, 2010; Shih, 2007) have broadly extended the aspects of the washback model. It is suggested that washback should be viewed inclusively, from micro aspects such as 'test, learner, and teaching', as well as macro level

impact aspects of ‘culture, family and community’ (Saville, 2010, p.4). With regard to washback on learning, the surface of washback (i.e. the contents and methods to learn by test influence) has been the major focus (e.g. Cheng, 1998; Qi, 2004; 2005), but the attention on underlying factors (i.e. the influence of students’ beliefs and sociocultural context) shaping such washback is growing (e.g. Tsang & Isaacs, 2022; Xie and Andrews, 2012). As Shih (2007) proposed, students’ extrinsic factors (e.g. friends, family, school) and intrinsic factors (e.g. students’ interests and beliefs) are nested within their learning actions. Thus, comparing these sociocultural contexts with washback is essential to identify how they mediate students’ learning (e.g. Tsang & Isaacs, 2022).

The present study analyses the relationship between washback on learning and mediating factors (i.e. sociocultural context), following previous studies (e.g. Booth, 2012; Shih, 2007; Tsang & Isaacs, 2022; Zhan & Andrews, 2014) in the high-stakes university entrance test in Hong Kong, the Hong Kong Diploma of Secondary Education Examination (HKDSE). HKDSE became the sole high-stakes test for secondary school students in 2012, replacing two former public examinations (HKALE & HKCEE). English language is one of the core subjects in the HKDSE assessment and has a major impact on stakeholders (e.g. students, teachers, and family) in washback literacy. While previous research has identified washback in the two former and replaced examinations (e.g. Qian, 2008), the HKDSE English language has received limited research focus. This study analyses how mediating factors (e.g. family, classmates, students’ beliefs) influence students’ test preparation by exploring the washback of speaking assessment (Paper 4) in the HKDSE.

Literature review

Washback on learning

Alderson and Wall (1993) proposed the first preliminary washback model on how testing influences teaching and learning, drawing from the empirical studies over the course of three decades. Before and during this period, teaching was the sole focus of washback studies (e.g. Andrews, 1994; Lam, 1994; Watanabe, 1996), while studies focusing on students were scarce (e.g. Tsagari, 2007a, 2007b; Watanabe, 1992). In an inquiry on validity in washback (Messick, 1996), Bailey’s (1996) ‘basic model of washback’ provided an engaging perspective founded on the micro aspects of washback, which shifted the focus of washback studies from teaching to learning, including out-of-class learning. Alderson and Hamp-Lyons (1996) student-centred study was the first to provide empirical evidence of students’ reliance on test material implying teacher influence of introducing test material to students. This study, among others (e.g. Andrews et al., 2002; Green, 2006) provided insight for later studies (e.g. Shih, 2007; Xie & Andrews, 2012) which focused on sociocultural factors shaping students’ learning choice.

Students’ learning is considerably prompted by the test; however, out-of-class learning methods can be varied and can occasionally be non-test-specific (e.g. Pan, 2014; Stoneman, 2006). Washback on learning was considered to be test oriented. Cheng (1997; 1998) provided evidence of students’ preference for test-specific activities over other activities in the classroom environment. This test-specific practice focus was confirmed by subsequent literature in Hong Kong (e.g. Stoneman, 2006; Tsang & Isaacs, 2022; Qian, 2008) and across Asia (e.g. Watanabe, 2004; Xie & Andrews, 2012). However, the use

of non-test-specific methods has been observed concurrently with test-specific methods in out-of-class learning (Stoneman, 2006; Pan, 2014; Xie, 2013). This includes general daily practices (e.g. watching movies and listening to English songs), online content (e.g. learning English using software), and interaction with native speakers. Students' beliefs (e.g. non-test-specific methods can help them improve), time constraints (e.g. using test-specific methods for short-term improvement), and opinions from others (e.g. following the learning method of their friends) are various factors that determine the degree of (non-)test-specific and method adoption (Baba, 2019; Green, 2007; Tsang & Isaacs, 2022; Pan, 2014). In addition, the growing use of out-of-class online learning (Zhan and Andrews, 2014) has been observed with methods such as test-oriented (e.g. test preparation websites), entertainment content (Stoneman, 2006), and game-play contents (Sylvén & Sundqvist, 2012). The variety of out-of-classroom learning can make it difficult to determine the nature of positive or negative washback (i.e. consider as good or bad learning methods); factors such as stakeholders' (e.g. students, test-makers) intention toward these learning methods (Green, 2007), long- and short-term learning goals (Maxmudjonova, 2023), and internalisation and results of various methods (Andrews et al., 2002) need to be examined. This study embraces the term 'observable washback' (Tsang & Isaacs, 2022) as the primary focus to describe learning actions, as positive or negative notions, should be examined carefully and used only after thorough investigation.

Impact and sociocultural washback

Washback on learning is not only affected by the test itself but also driven by the agents and power with and without the students (Andrews et al., 2002; Cheng et al., 2011; Chik & Besser, 2011; Tsang & Isaacs, 2022). The underlying interaction between multiple aspects other than the test itself is often associated with 'impact' research (e.g. Booth, 2012; Shih, 2007; Shohamy, 2007; Stoneman, 2006; Tsang, 2017; Xie & Andrews, 2012). The impact research explores mediating factors from various stakeholders (e.g. students, teachers, and family) or involved elements (e.g. test, environment, and society). Shih (2007, 2010) proposed a sociocultural washback model of students' learning based on various types of mediating factors. His model suggests that students' learning is not only affected by the test but also determined by the extrinsic and intrinsic sociocultural factors. His study on existing tests in Taiwan, using classroom observations and interviews, found a small but varied washback in parts of the school. Utilising various types of washback, he proposed a model with external (e.g. family and teachers), intrinsic (e.g. language proficiency and motivation), and test (e.g. stakes and content) factors to explain his findings. According to the results, these factors interact differently with the washback effect. This provided insight into how washback is sharpened by these mediating factors rather than solely observing stakeholders' actions.

An awareness of these mediating factors has led others to focus on selected factors. (e.g. Cheng & Deluca, 2011, Fox and Cheng, 2007; Xie & Andrews, 2012). Fox and Cheng's (2007) study, utilising class observations and interviews, investigated test preparation for the Ontario Secondary School Literacy Test in Canada. Students' intrinsic factors were the reason for lower self-perceived language proficiency of English as a second language (ESL) students compared with their native peers. Alternatively, Xie and Andrews (2012) examined other aspects of intrinsic factors in the Chinese College

Entrance Test (CET). Their study proposed four intrinsic factors from expectancy-value theory (Jacob & Eccles, 2000) and two related aspects (test design, test use) and their influence on students' test preparation. These studies, among others, indicated possible intrinsic factors such as language proficiency (e.g., Pan, 2014) and test-taking experiences (e.g. Cheng & Deluca, 2011; Stoneman, 2006) for future research.

In addition to intrinsic factors, Green's (2006) impact study on International English Language Testing System (IELTS) in a Chinese setting revealed extrinsic factors at the school level. This study was quantitative and revealed test-specific preparation methods adopted by students through rounds of questionnaires. He concluded that teachers' expectations were behind students' adoption of test-specific preparation methods. Teachers' language assessment literacy (Tsiplakou & Tsagari, 2023), teachers' viewpoints toward the test (Tsagari, 2020), their feedbacks given to students (Khan et al., 2023; Samaie & Valizadeh, 2023), and score-related pressure given to teachers (Ali & Hamid, 2023) can reciprocally determine students' washback. Chik and Besser's (2011) study on preparation for an international language test in Hong Kong exhibited other extrinsic factors. Multiple phases of data collection in several settings (e.g. interviews, classrooms, and online observation) suggest that school environment, cram schools, and the Hong Kong media influence Hong Kong learners to some degree. These studies, among others, indicated possible extrinsic factors such as peer information (e.g. test score) from private tutoring (Allen, 2023), peer evaluation (e.g. Black et al, 2003; Cowie, 2005), teacher belief and evaluation (e.g. Gu & Lyu, 2023; Khan et al., 2023), and social language attitude (Shum et al., 2023) for future research.

Similar to Shih (2007), Tsang and Issacs (2022) examined the washback effect in a sociocultural approach that focused on conclusive factors from stakeholders in and beyond the classroom. They found negative washback in learning from the introduction of the new test design, the graded approach in the HKDSE English writing test. They concluded that learners' perceived washback effects were affected by mediating factors from multiple stakeholders (e.g. teachers, classmates, and family members). For example, learner test preparation using intensive paper and pencil drills was linked to teacher evaluation and learner's knowledge of the test. All types of test preparations were affected by at least one of the mediating factors, which affirmed the influence of intrinsic and extrinsic factors on HKDSE. He concluded that 'washback on learning is fundamentally a construct driven by an array of intertwining forces' (p.59).

HKDSE-English Language Paper 4

The Hong Kong Diploma of Secondary Education Examination (HKDSE) was introduced in 2012 as a high-stake and standard reference test for senior secondary schools conducted annually in Hong Kong. Of the three compulsory subjects offered in the HKDSE, English (HKDSE-English) is administered for assessing four objectives: reading, writing, listening, and speaking. The assessment scheme (a total of 100 points) assigns the following weights, 20% to reading, 25% to writing, 30% to listening and integrated skills, 10% to speaking (0% in the 2020 HKDSE), and 15% to school-based assessment (25% in the 2020 HKDSE owing to the cancellation of the speaking test) (Hong Kong Examinations and Assessment Authority, 2020). The focus of this study is HKDSE-English Language Paper 4, which consists of two parts: (A) group interaction (8 min per

group of four candidates) and (B) individual response (1 min per candidate). In group interaction, the candidate is required to discuss an assigned topic (10-min preparation time given) with three other test-takers. Little instruction is given during the test, such that the discussion is unrestricted and extemporaneous as 'you will be... completely responsible for the progress of the discussion' (Hong Kong Examinations and Assessment Authority, 2019, p.1). The examiner poses questions to individual candidates for which no time or preparation is given before answering. This unique and close to real-life method in a high-stakes test is believed to have sound validity and facilitates positive learning and teaching (Qian, 2008). Furthermore, school-based assessment (SBA) that aligned with the test was used for continuous evaluation and improve school learning (Davison & Hamp-Lyons, 2009; Yu, 2010).

To the best of the researcher's knowledge, no research has been conducted on the washback on learning from HKDSE-English Language Paper 4. Several researchers have examined previous speaking tests before the new reforms (Andrews, 1994; 2002; Cheng, 1998, 1999; Qian, 2008) or related school assessment section (e.g., Yu, 2010). These have neither focused on teaching nor have examined these tests inclusively with other test areas (i.e. reading and writing). The focus on HKDSE has been scarce for a decade; the latest research has investigated other sections (i.e. writing section) of the test (Tsang, 2017; Tsang & Isaacs, 2022). Thus, there is a significant need for research that reaffirms the washback on learning from the HKDSE-English Language Paper 4 and further examines the mediating factors underlying it.

Washback studies, conducted in Hong Kong (e.g. Chik & Besser, 2011; Stoneman, 2006; Tsang, 2017), provided insights into test-specific and general preparation methods used by local students. Tsang and Issacs (2022) used the term 'observable washback' and classified it into two categories: explicit ways of practising (target skills) and implicit ways of practising (target skills). Similar to test-specific and general preparation, explicit ways of practising indicate that learner action or preparation is directly centred, whereas implicit ways of practising indicate that learner action or preparation is indirectly focused on the test. The term 'observable washback' is adopted as a neutral category term, aside from positive–negative trichotomy.

In summary, Shih (2007) proposed a sociocultural washback model that provided the fundamental mediating factor theory in this research. Previous washback studies showed evidence that learner perspective, attitude, and their surrounding affect washback equally with the test introduction. Thus, learning toward test does not stand alone but is sharpened by mediating factors. In other words, beyond the test itself, mediating factors for individual learner and society should be examined. This research seeks to examine the relationship between washback and mediating factors (Shih, 2007) empirically then further affirm the previous findings in the HKDSE environment (Tsang & Isaacs, 2022). Under Shih's model, Tsang and Isaacs's study using simultaneous multiple regression (SMR) provided a separated but apparent relationship of each types of explicit test preparation to mediating factors. While their result provides meaningful insights into mediating factors, two empirical gaps remain: (1) what types of students exist in the current test environment and how they choose from different test preparation (i.e. assuming students would not only use exclusively one type but also multiplied types of preparation during the test period), reciprocally affected by mediating factors, are unknown. (2).

Unlike other studies in Hong Kong (e.g. Stoneman, 2006), general preparation methods were not examined. This raises the question of how students' implicit learning functions in HKDSE. By filling the research gaps, a broader and student-oriented perspective of washback and mediating factors can be expected.

Research questions

To address the lack of attention into HKDSE and affirm the findings from previous studies, the current study is based on a conceptual replication approach (Tsang, 2017; Tsang & Isaacs, 2022). The following research questions are framed:

1. What are the observable washback effects of students (test-takers of HKDSE English Language Paper 4) when preparing for the test outside the classroom?
2. How do the mediating factors from previous findings (Tsang, 2017; Tsang & Isaacs, 2022) shape students perceived observable washback effects?

Methods

Participants

The survey data was collected in June, 2020. Two-hundred and thirty-five students from a local secondary school were selected, and using purposeful sampling due to the nature of the school is relevance to the research purpose. The school was chosen as it is government-aided institution and also uses Chinese as the medium of instruction. This is representative for identity of the majority of HKDSE test-takers characteristic, where approximately 71% of schools are Chinese medium-of-instruction (CMI) government aided (Education Bureau, 2022, p.63). Two-hundred and seventeen completed questionnaires were received (39.3% males and 59.7% females) with a response rate of 92.3%. For eliminate false data skewing the results, eighteen questionnaires with identical answers (the same responses for all items) were excluded. Table 1 shows the summary of one-hundred and ninety-nine participants (36.2% Form 4, 27.1% Form 5, and 36.7 Form 6). Prior to the main study, a pilot study ($n=32$) was administered in March 2020. These students were also included in the main survey.

Research instrument

A 5-point Likert scale questionnaire with 52 items was administered (see Appendix A). For test preparation methods, 26 items measuring various extracurricular learning

Table 1 Demographics of participants

		Gender		First language		Study abroad experience
High school year	<i>n</i>	Male	Female	Chinese	English	
Year 1 (Secondary 4)	72 ^{*1}	28	42	70	2	3
Year 2 (Secondary 5)	54	20	34	54	0	9
Year 3 (Secondary 6)	73 ^{*2}	24	48	72 ^{*3}	0	4

^{*} 1, gender unknown = 2; ^{*}2, gender unknown = 1; ^{*}3, first language others = 1

activities were selected from multiple studies (Pan, 2014; Stoneman 2005; Sylvén & Sundqvist, 2012; Tsang, 2017; Tsang & Isaacs, 2022). These included both test-specific (i.e. explicit ways of practising) and test-unspecific (i.e. implicit ways of practising) learning activities. For sociocultural factors, 26 items were selected from the questionnaire in Tsang and Issacs (2022) study to measure mediating factors.

For example, the use of tutorial schools from Pan's (2009) and Tsang and Issacs's (2022) findings, preparation through communication from Stoneman's (2005) finding, or the use of entrainment from multiple findings (Pan, 2009; Stoneman, 2005; Sylvén & Sundqvist, 2012) were included. The siblings mediating factors were excluded from the study because of the unfitting factor analysis in our pilot state. This instrument aims to reveal the relationship between observable washback (i.e. test-specific and test-unspecific extracurricular activities) and mediating factors (Shih, 2007; Tsang & Isaacs, 2022).

Data collection and analysis

A pilot test was conducted with 32 Form 5 student participants in late March 2020 to ensure reliability, item readiness, and wording appropriateness. The overall result indicated high reliability (Cronbach's $\alpha = 0.955$) and acceptable item readiness (three comments about item readiness in 32 questionnaires).

The main survey was conducted with 235 student participants through Google Forms in mid-May, 2020. This was an intense period for test preparation, as the cancelled 2020 speaking test (Hong Kong Examinations and Assessment Authority, 2020) was originally arranged to take place in May 2020. Owing to the relatively larger number of participants and the cancelled test, the data was further collected online for 2 weeks until late May. The adjusted questionnaires in the main survey were provided in both Chinese and English languages and included 52 items.

Data analysis was performed using IBM SPSS Statistics version 25 (SPSS 25). Two separate sets of exploratory factor analyses (EFA) were conducted: (1) observable washback effects and (2) mediating factors utilising the principal axis factoring method and the Promax rotation method. Sampling adequacy was confirmed using the Kaiser–Meyer–Olkin (KMO) index (both > 0.5) and Bartlett's sphericity test (both < 0.05). Item extractions were based on Kaiser's criterion of retaining factors with eigenvalues greater than 1, given that each factor accounts for at least 7% of the total variance (Fields, 2013; Stoneman, 2006, Tsang & Isaacs, 2022). Consequently, six clusters of observable washback and five clusters of mediating factors were identified. The results from the two sets of EFA were further compared using cluster analysis. This study aims to statistically analyse the relationships between the variables.

Results

Research question 1: What are the observable washback effects of students (test-takers of HKDSE English Language Paper 4) when preparing for the test outside the classroom?

To answer RQ1, exploratory factor analysis (EFA) with 26 extracurricular test preparation items was conducted. Sampling adequacy was achieved with a KMO value of 0.850 and significant fit in Bartlett's test of sphericity (chi-square = 3471.52, $p < 0.001$). Factor analyses obtained clusters based on Kaiser's criterion for eigenvalues > 1 , and six clusters were extracted, with 70.61% of the variance explained. Table 1 shows the results

of exploratory factor analysis (EFA) with 26 items on the extracurricular test: (1) integrated practising English daily and online content ($\alpha=0.88$), (2) practising English with test-specific material and drilling ($\alpha=0.87$), (3) practising English with communicative purposes ($\alpha=0.84$), (4) practising English in digital gaming ($\alpha=0.94$), (5) practising English in tutorial classes ($\alpha=0.86$), and (6) practising English in entertainment content ($\alpha=0.79$).

Five of the six factors were organised into two groups of observable washback: implicit and explicit ways of practising. The first factor, containing both test specific (e.g. using online test preparation website) and test unspecific (e.g. reading English novels, listening to radio programmes), was listed as 'integrated practice in daily and online contents', as its nature could not be clearly defined. Table 2 shows that the first group represented the implicit way of practising the target language and included three out of six factors (Factors 3, 4, and 6). The other group represents the explicit ways of practising the target language and included the last two factors out of six (Factors 2 and 5). Factor 1 was placed in the middle because of its integrated nature.

Further analysis was conducted on six factors of observable washback to compare the differences. Table 3 presents descriptive statistics for the six factors with Factor 1 as the average. The contents included multiple learning methods (e.g. finding test-related information online, listening to radio programmes, reading newspapers) that correspond to the parallel use of test-related and non-test-related learning. Factor 6, as an implicit factor (practising in entertainment content), was rated as the highest, compared to Factor 2 (practising with test-specific material and drilling), both in the value of mean (mean = 3.87 > 3.55) and the rank test ($Z = -0.4.83$, $p < 0.000$). This suggests that in HKDSE-English Language Paper 4, students favoured implicit practice methods more than traditional test preparation methods. The comparison of differences in observable washback provides insight into students' preferences, as the main purpose of the results was cross-analysis with mediating factors.

Research question 2: How do the mediating factors from previous findings (Tsang, 2017; Tsang & Isaacs, 2022) shape students perceived observable washback effects?

The EFA was conducted with 26 mediating factors (Tsang, 2017; Tsang & Isaacs, 2022). The items in the mediating factor effects were adopted from a previous study in an identical environment. Similar to RQ1, an EFA using the extraction method of principal axis analysis with Promax rotation was performed on the 26 items. Sampling adequacy was achieved based on a KMO value of 0.899 and a significant fit on Bartlett's test of sphericity (chi-square = 4311.68, $p < 0.000$) in the preliminary analysis. The factor analyses resulted in five clusters which were extracted based on Kaiser's criterion of eigenvalues over 1, with 72.08% of the variance explained. Table 4 presents the factor loadings after the Promax rotation. Based on the nature of clusters and referring to patterns from the Tsang and Issacs (2022) study, five factors were extracted as follows: (1) family and tutorial influence ($\alpha = 0.94$), (2) interest in the language ($\alpha = 0.89$), (3) evaluation made by teachers ($\alpha = 0.87$), (4) language proficiency ($\alpha = 0.81$), and (5) peer influence ($\alpha = 0.92$). The five factors hold a pattern similar to Tsang and Issacs' findings (eight factors were found); however, some mediating

Table 2 Factor analysis of students' observable washback

Pattern matrix ^a						
Item	Factor loading					
	1	2	3	4	5	6
Factor 1: Integrated practising English daily and online contents						
P06. Reading English novels and magazines	0.82	-.03	.04	0.13	-.12	-.03
P07. Writing in English (e.g. dairy, memo, story/novel)	0.75	-.08	0.26	-.09	-.05	-.08
P03. Listening to English radio programmes	0.70	-.06	-.05	-.06	-.01	-.06
P05. Reading English newspapers	0.67	0.17	.01	-.12	-.16	0.17
P16. Finding information about the test on online forum	0.65	0.18	-.06	.01	.05	0.18
P17. Spending time on test preparation-specific websites	0.53	0.16	.00	.00	0.20	0.16
P12. Surfing or enrolling in online forum for daily use	0.41	-.09	0.15	0.30	0.19	-.09
P11. Surfing English websites	0.39	-.07	.01	0.39	0.11	-.07
Factor 2: Practising English with test-specific material and drilling						
P23. Drilling of vocabulary materials (e.g. word cards, textbook)	.05	0.87	.03	-.08	-.06	-.07
P24. Memorising information about the test and getting used to it (e.g. test format, timing, question types)	.02	0.80	.01	0.10	-.24	.02
P18. Doing past exam papers of DSE English Paper 4	0.19	0.69	-.18	.00	0.11	-.11
P20. Enrolling in or organising a mock examination	-.025	0.64	0.13	.06	0.13	0.11
P22. Drilling of grammar materials (e.g. class notes, textbook)	0.29	0.63	-.10	-.05	0.10	-.01
P21. Enrolling in or organising a study group with class-mates	-.020	0.53	0.21	-.14	0.27	0.18
Factor 3: Practising English with communicative purposes						
P08. Writing to people in English (e.g. SNS, messages, letters)	0.21	-.16	0.83	-.09	.06	-.01
P09. Seeking opportunities to speak with native speakers of English	0.14	0.21	0.67	0.17	-.20	-.10
P15. Using SNS with others in English (e.g. WhatsApp, Skype, FaceTime)	-.014	0.11	0.65	0.15	.01	-.02
P10. Speaking English with friends and family members for practice purposes	0.21	-.07	0.58	-.21	0.15	0.10
Factor 4: Practising English in digital gaming						
P13. Playing online games that use English to communicate	-.04	-.02	.01	0.89	.05	.04
P14. Playing single player games that require reading in English	-.05	.02	.01	0.89	.00	0.12
Factor 5: Practising English in tutorial classes						
P25. Finding and enrolling in tutorial classes that focus on DSE English Paper 4	-.06	-.05	.01	.02	0.98	.00
P26. Continue to enrol in a tutorial class when the tutor can improve my preparation	-.05	0.11	-.01	.08	0.80	-.03
P19. Drilling via a test preparation specific commercial book, CD, or DVD	0.19	0.30	.01	-.02	0.43	-.08
Factor 6: Practising English in entertainment contents						
P01. Watching English movies videos	-.01	.02	.01	0.11	.03	0.80
P02. Watching English TV programmes and dramas	0.29	-.14	-.14	.01	0.10	0.69
P04. Listening to English songs	.04	0.11	.08	.05	-.18	0.65

N = 199. The extraction method was principal axis factoring with an oblique (Promax with Kaiser normalisation) rotation. Factor loadings above 0.30 within individual factor are in bold

^a rotation converged in seven iterations

Table 3 Factor groups of the six observable forms of washback

Implicit ways of practising the target language	Explicit ways of practising the target language
3) Practising with communicative purposes	2) Practising with test-specific material and drilling
4) Practising in digital gaming	5) Practising in tutorial classes
6) Practising in entertainment contents	
1) Integrated practice in daily contents and online	

Table 4 Descriptive statistics of the six factors of observable washback

Factor	Mean	Std. deviation
1) Integrated practice in daily contents and online	3.32	0.75
2) Practising with test-specific material and drilling	3.55	0.74
3) Practising with communicative purposes	3.15	0.89
4) Practising in digital gaming	3.57	1.04
5) Practising in tutorial classes	3.13	0.94
6) Practising in entertainment contents	3.87	0.77

factors were either combined with other factors (e.g. family and tutorial school) or had a weak loading to be independent in EFA (e.g. society influence). The findings from the EFA provided the foundation for research question 2. Combining the analyses with findings from research question 2 (i.e. six types of observable washback), correlation was performed using cluster analysis (Table 5).

Cluster analysis was conducted to identify the patterns in factors of students' learning and mediating. Six types of students' washback effect and five categories of mediating factors were obtained based on the clustering. As shown in Fig. 1, the 11 factors are presented in a left-to-right order similar to the previous corresponding order. The results indicate four types of HKDSE test-takers under both washback and mediating factor. The mean of the group provides the degree of the student's dedication while preparing for the test.

The four clusters are arranged in a descending order based on the number of participants and are characterised as follows:

1. Cluster 1 ($n=81$) shows balanced means for all factors, exhibiting an above average score. The balanced mean value of learning methods and mediating factors indicates no strong preference or influence. The less significant family mediating factor is similar for all the clusters.
2. Cluster 2 ($n=71$) shows the lowest average among all clusters. It is characterised by higher preparation using test-specific material and entertainment, corresponding to a higher mediating factor in language interest and self-perceived language proficiency. The results suggest that students are less dedicated to test preparation, and their learning preference is in using test-specific and entertainment content. Strong intrinsic mediating factors suggest that students decide their learning choices.
3. Cluster 3 ($n=30$) is characterised by higher game-based and entertainment test preparation, a higher mediating factor in language interest, and self-perceived lan-

Table 5 Factor analysis of mediating factors

Pattern matrix ^a					
Item	Factor				
	1	2	3	4	5
Factor 1: Family and tutorial influence					
S20. Tutorial schools' advice on preparatory work	0.96	.07	-.03	-.014	.03
S19. Tutorial schools' advertisements	0.94	0.10	-.011	-.013	.06
S21. Test-taking skills learned from tutorial school	0.90	.05	-.04	-.013	.06
S24. Reports published by the media	0.69	.00	0.32	.09	-.027
S22. Information for preparatory work on websites	0.69	-.02	0.18	-.02	.07
S17. Family's advice on preparatory work	0.66	-.010	-.07	0.23	0.25
S23. Posts about preparatory work on online forums	0.56	-.02	0.39	.01	-.011
S18. Family's expectations	0.48	-.016	.03	0.35	0.22
Factor 2: Interest in the language					
S04. The way I use English	.03	0.85	-.02	.00	.06
S05. The value of English as a language	0.14	0.77	.00	-.04	-.01
S06. My prioritisation of English subject	0.16	0.72	-.016	0.10	-.04
S03. My ambitions	-.017	0.71	0.30	-.012	0.11
S02. My learning goal	-.027	0.66	0.18	0.10	.08
S07. My interest in the English subject compared to other subjects	0.18	0.64	-.018	0.30	-.020
Factor 3: Evaluation made by teachers					
S15. Teachers' assessment of my English language ability	-.02	-.06	0.87	.08	.02
S16. Comparisons made by teachers (e.g. scores with previous years or between classmates)	0.13	-.09	0.70	.03	.09
S14. Teachers' expectations	0.13	.05	0.65	-.02	0.12
S25. The examination-oriented culture in Hong Kong	0.13	-.01	0.63	.09	-.08
S13. Teachers' advice on preparatory work	.05	0.19	0.60	-.04	0.13
S26. The importance of the HKDSE English exam in modern Hong Kong society	-.017	.01	0.30	0.25	0.21
Factor 4: Language proficiency					
S09. My English score on school English examinations	-.012	.09	0.16	0.76	.04
S08. My English language ability	-.08	0.18	0.12	0.71	-.011
S10. My performance in English classroom activities	.08	0.20	-.03	0.53	0.13
Factor 5: Peer influence					
S12. Classmates' score on the school English examination	0.19	-.01	0.37	-.08	0.62
S11. Classmates' preparatory work	0.19	-.01	0.20	.07	0.59

N = 199. The extraction method was principal axis factoring with an oblique (Promax with Kaiser normalisation) rotation. Factor loadings above 0.30 in individual factor are in bold. ^a rotation converged in seven iterations

guage proficiency. The cluster also has a strong lean, with the lowest tutorial test-preparation and mediating factors in family and peers. The results suggest that students prefer enjoyable daily content and consider it useful for their English learning and test preparation. These results are similar to Cluster 2 in mediating factors. Such implicit learning choices are decided by intrinsic factors, such as students' language interest and proficiency.

- Cluster 4 ($n=17$) is characterised by higher tutorial and test-specific material test-preparation, higher mediating factors, with peer influence being the highest mediating factor. The cluster has relatively lower learning use in game-based and entertainment. The results suggest that the student group is most dedicated to test preparation. The noticeable use of explicit methods (i.e. test-specific material and

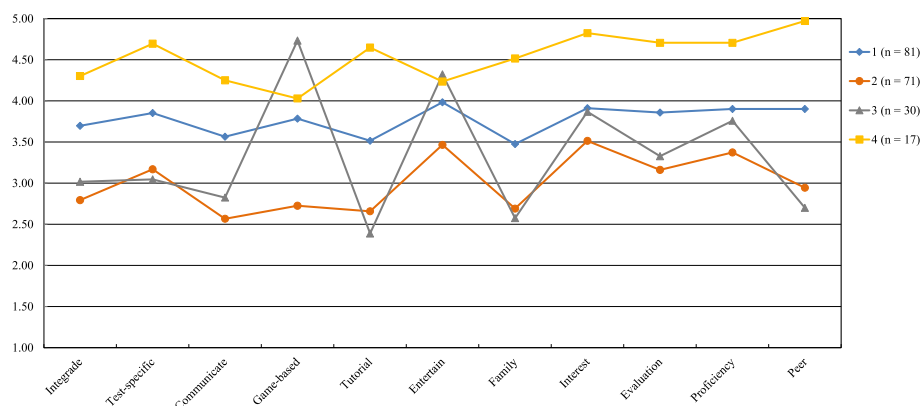


Fig. 1 Cluster analysis of six types of washback and five types of mediating factor

tutorial school) suggests the students group prefer the ‘traditional way’ in test preparation. Apart from explicit methods, all other learning methods were favoured to some degree. Thus, test-dedicated students use all learning methods in their learning strategy. The extremely strong peer influence suggests that classmates’ opinions affect students’ learning with considerable involvement of study groups.

Discussion

Regarding RQ1, six types of learning method clusters were adopted by the students. Students engaged in autonomous learning using non-test-specific methods (i.e. focus on language and less relevant to test) slightly more than test-specific methods. One of the favoured method clusters was of practising using entertainment content and digital gaming.

Regarding RQ1, learning with test-specific material was favoured by the most and the least dedicated group. Tutorial schools were the least favoured method.

Regarding RQ2, evidence was provided by four student groups of test-specific (explicit) methods being driven by extrinsic factors. Strong peer influence was found in most dedicated and test-specific groups in contrast to groups that preferred non-test-specific methods. The use of test-specific methods was considerably driven by extrinsic mediating factors.

Regarding RQ2, non-test-specific (implicit) methods were driven by intrinsic factors. Groups which used entertainment and game-based methods had high correlation with language interest and proficiency. The extrinsic factors had lesser influence on non-test-specific groups, such as family, teachers, and peers.

This study aims to reveal the underlying sociocultural factors that affect students’ observable washback following the theoretical works of Tsang and Issacs (2022). The results exhibited two major correlations between the students’ observable washback (implicit/explicit) and sociocultural factors. First, students who preferred test-specific methods considered that other opinions (e.g. peers, family, and teachers) affected their choice. This indicates that the influence of extrinsic mediating factors can lead to explicit learning. Second, students who preferred non-test-specific methods made a choice based on their own individual factors. Thus, the use of implicit learning is driven by

intrinsic factors and vice versa. The results suggest that students' learning choices are influenced not only by the test but also the sociocultural factors which greatly affect how students learn.

Tsang and Issacs (2022) had prioritized on explaining the reasons for certain learning actions. This study further distinguishes which student groups could be predicted with mediating factors, in the HKDSE English environment: the most dedicated 'top students' groups (Cluster 4), the average group (Cluster 1), the self-driven entertainment group (Cluster 3), and the least dedicated 'self-driven but do what is necessary' group (Cluster 2).

The two self-driven groups (Clusters 2 and 3) show a major influence of sociocultural factors, specifically in their language interest and proficiency. The use of entertainment content as a learning method was found in studies in Hong Kong decades ago (Cheng et al., 2011; Stoneman, 2006). Similarly, the focus on test-specific material used was considered an effective score gaining method in the Asia environment (Pan, 2014; Stoneman, 2006). Interestingly, the use of entertainment and test-specific methods for learning in Cluster 2 is different from the others. Thus, lesser-dedicated students select their strategy practically to achieve what is necessary, one 'enjoyable' way to learn, and one 'practical' way to scoring gaining. This result is consistent with Stoneman (2006) that students preferred such non-test-specific methods and considered them 'enjoyable'. In the manner of improving the test scores as their short-term (immediate) learning goal, students tend to employ test-specific learning strategies (Dong, 2023; Maxmudjonova, 2023). On the contrary, digital gaming as implicit in English learning has not yet been reported but is emerging from the European environment (e.g. Sylvén & Sundqvist, 2012). The high preference for game-based learning methods reflects that a growing number of students selected practising from games that required English skills as their learning strategy. The game-based learning use in Cluster 3 provides enjoyable ways to learn; this may help serve learners' long-term goal (Stoneman, 2006). The differences in game type and content can vary the learning quality significantly (Gee, 2007). Students consider it helpful in their test preparation, but the games played should be carefully analysed.

In terms of mediating factors, the two groups prefer entrainment or gaming content based on their language interest and proficiency. In Tsang and Issacs's (2022) study, these two factors were associated with activities outside the classroom. This suggests that such activities and learning practices of the two groups are autonomous. This preference for autonomous practising, unlike the use of traditional test requires preparation (e.g. the use of test-specific material and tutorial school), is not unique to HKDSE but matches previous washback studies (e.g. Pan, 2014; Stoneman, 2006). Based on the nature of the cluster, such autonomous practice is for learners who are associated with intrinsic factors. Thus, to achieve autonomous learning as washback, increasing students' strong language interest and understanding their own proficiency level are essential. On the other hand, the result reveals that both groups exhibit lower extrinsic factors, such as family and teachers influence. It thus seems that family and teachers influence are 'negative', as they somehow against students choosing the 'positive' autonomous practising. It is important to note that such tendency may stem from the discrepancy between the entrainment-gaming contents learning and 'traditional learning', while their influence does not necessary against the novel learning approach. For example, teachers influence

can improve students' learning, such as encourage peer/self-evaluation and give constructive feedback (Khan et al., 2023; Samaie & Valizadeh, 2023) and assume teachers to crucial role of monitoring students' entrainment-gaming learning strategies (Sylvén and Sundqvist, 2012).

The most dedicated student group (Cluster 1) and their noticeable use of test-specific materials and tutorial schools show preferences for traditional methods (Stoneman, 2006; Zhan & Andrews, 2014). Explicit learning in test-oriented drilling, material, and tutorial schools was the most common in Hong Kong and Asian environments (e.g. Allen, 2016, 2023; Stoneman, 2006; Tsang & Isaacs, 2022). This indicates that students consider traditional preparation to be the most useful method for HKDSE. Unlike other groups, dedicated test-takers use implicit and explicit learning inclusively, suggesting they consider both improving their test preparation. The inclusive use of multiple methods, including intergrade learning and game based, reflects the adoption of online learning, as useful test website and enrolling in school are commonly provided to students (Pan, 2014). The result of Cluster 4 resembles Pan's findings (2014), as students adopted both implicit and explicit learning during their initial years of high school, while during final years (with higher test intensity), highly motivated (dedicated) students adopted more test-specific methods for gaining scores.

In terms of mediating factors, the most dedicated group is affected by language interest and proficiency but largely by extrinsic factors such as peer influence. Strong language interest is common among dedicated test-takers. In contrast, extremely strong peer influence suggests that students share their preferences or even engage in dedicated group to seek the most sophisticated learning methods, thus affecting their preference (i.e. test-oriented material and tutoring). This strong peer influence on extrinsic learning is consistent with Tsang and Issacs (2022); also, student's preference on relying on peer helps and ideas was suggested (e.g., Black et al, 2003; Cowie, 2005). Furthermore, the result aligns with Allen (2023) that private tutoring in Japan provides academic performance information relative to their peers. This interplay between the uses of private tutoring, peer influence, and tutoring influence exemplifies that mediating factors affect beyond students' learning choices, with the reciprocal effect can potentially amplify their own influences through the learning process.

Thus, we conclude that our results have been demonstrating the critical role of implicit and explicit mediating factors in determining observable washback, together with the test itself. The widely held belief that the test alone determines its washback is challenged by our finding. We have revealed that mediating factors carry greater influence over students than the test itself, especially if a period has elapsed since the introduction of the test. Our study has illuminated that these washback and mediating factors engage in a reciprocal relationship, rather than a casual one. Such as, students who enjoy entertainment content will have a higher language interest which may lead to demand for more entertainment content or extend implicit learning. These findings confirm the influence of extrinsic and intrinsic mediating factors to students' learning choices. Importantly, our study advances the understanding of the existing students characteristics by offering a model of where learning and sociocultural contexts engages in high-stakes test environment. This interactive dynamics of mediating factors and student learnings suggests careful analysis of stakeholders elements such

as belief and opinions, and their surrounding is needed to achieve positive washback. Future research should examine the complex structure of such reciprocal relationship. To summarise, by redefining the washback concept within the sociocultural context, our study adds to the growing body of research (e.g. Cheng et al., 2015; Shih, 2007; Tsang & Isaacs, 2022) which acknowledges the complex interplay between testing, learning, and social influence. Our findings provide a novel and student-oriented perspective for facilitating positive washback in HKDSE English.

Conclusions

This study has several implications in the field of washback and for Hong Kong educators. Initially, using the classification of observable washback, which objectively observes the nature of students' test preparation, it contributes to revealing the notion of washback by (a) addressing the lack of studies that investigate HKDSE-English Language Paper 4 after the revision of 2012 and (b) addressing the relationship between washback and mediating factors in the student groups.

This research provides new insights on out-of-class learning methods among current Hong Kong secondary students. The two major results of the study are high preference for implicit (non-test related) methods adopted by majority of students and strong peer influence (extrinsic factors) associated with explicit methods. In the school environment, implicit methods have been extensively adopted by students. This indicates that teachers or course makers can adopt entertainment-related activities for test preparation. New learning methods for test preparation such as digital gaming may extend the possibilities of future washback studies. However, the connection between explicit methods and peer influence should consider how they shape students' inner groups. Future research can draw conclusions on what constitutes 'negative washback' from test-specific methods because the group discussion test method is considered closer to daily speaking than other tests. For students to understand the test, test-related preparations should match to daily speaking, thus raising the question, 'Is studying for the test (explicit methods) equal to negative washback?' For defining these test-related preparations in a positive–negative dichotomy of washback, examining only 'what' students do (as conducted in the study) is not enough. In a type of learning action (e.g. drilling students on past paper), 'how' refers to how they prepare (e.g. is the procedure/material test-specific and authentic to daily content) and how they feel when they prepare (e.g. do they feel bored, anxious, or more motivated) and the 'why' behind the selection of a particular method (e.g. is it suggested by teacher/peer?). Further research that conclusively examines all facets is required. To facilitate learning of general English, results from clusters can be improved, which suggests that the use of implicit learning is linked to students themselves, while explicit learning is linked to extrinsic factors such as peers.

It is important to note that the HKDSE also includes a school-based assessment which shares a similar scoring system (Yu, 2010), and a lack of examination of SBA may not fully reveal the ongoing washback.

This study has three major limitations inherent to the quantitative research design, which limits the data use and generalisation. First, the data collection method of

using questionnaires limited the representation of the participants. As the questions were close ended, the students could not voice any opinions other than the selected items. There was also a lack of integration as these selected items were adopted from separate literature. Therefore, any correlations or relationships drawn across different sections of the study may not be valid or fully representative. Second, as a conceptual replication, the items of EFA and cluster analysis are partially adopted from previous research. Thus, the validity of the variables may not hold true as in the original research. Finally, the reciprocal relationship between washback and mediating factors requires a structural investigation. A thorough qualitative research involving interviews and continued observation should be conducted in the future to identify the progress and degrees of certain mediating factors influencing students' learning.

Future studies should employ a mixed-method approach to examine the washback phenomenon. The nature of 'negative' washback, to reveal multiple facets underlying learner actions, requires diversified research and a mixed research approach which should include data collection from interviews and class observations and not rely on a self-reported method. A follow-up research with a better and more sophisticated research design is required to confirm research washback finding for HKDSE-English Language Paper 4.

Abbreviations

CET	Chinese College Entrance Test
CMI	Chinese medium of instruction
ESL	English as a second language
EFA	Exploratory factor analysis
HKDSE	Hong Kong Diploma of Secondary Education Examination
HKALE	Hong Kong Advanced Level Examination
HKCEE	Hong Kong Certificate of Education Examination
IELTS	International English Language Testing System
SBA	School-based assessment

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40468-023-00250-4>.

Additional file 1: Appendix A. Research Instrument Item List (Provided on Google Forms).

Acknowledgements

The author gives special thanks to students and teachers who help or participate in this research from the local secondary school in Hong Kong.

Author's contributions

The research, data collection, data analysis, data interpretation, findings, and conclusion were solely done by the author. The author read and approved the final manuscript.

Funding

This work was supported via funding from JST SPRING (JPMJSP2150).

Availability of data and materials

The datasets generated and analysed during the current study are available in the OSFHOME repository, URL: https://osf.io/v65au/?view_only=c48ad7229d1e4303a9770459c27e6c8c.

Declarations

Competing interests

The author declares no competing interests.

Received: 28 April 2023 Accepted: 6 July 2023

Published online: 10 August 2023

References

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Alderson, C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing - LANG TEST.*, 13, 280–297. <https://doi.org/10.1177/026553229601300304>
- Andrews, S. (1994). The washback effect of examinations: Its impact upon curriculum innovation in English language teaching. *Curriculum Forum*, 4(1), 44–58.
- Andrews, S., Fullilove, J. & Wong, Y. (2002). Targeting washback – A case-study. *System*, 30(2), 207–223. [https://doi.org/10.1016/S0346-251X\(02\)00005-2](https://doi.org/10.1016/S0346-251X(02)00005-2).
- Allen, D. (2016). Investigating washback to the learner from the IELTS test in the Japanese tertiary context. *Language Testing in Asia*, 6 (7). <https://doi.org/10.1186/s40468-016-0030-z>
- Allen, D. (2023). English private tutoring and washback in Japan. *International Perspectives on English Private Tutoring: Theories, Practices, and Policies*. (pp. 37–54). Springer International Publishing.
- Ali, M. M., & Hamid, M. O. (2023). Teacher agency and washback: Insights from classrooms in rural Bangladesh. *ELT Journal*, 77(1), 52–61.
- Baba, S. (2019). How to produce beneficial washback effect by using high-stakes testing? Proposal from Educational Psychology. *JLTA Journal*, 22, 44–64. https://doi.org/10.20622/jltajournal.22.0_44
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13, 257–279.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning: Putting it into practice*. Open University Press.
- Booth, D. K. (2012). *Exploring the washback of the TOEIC in South Korea: A sociocultural perspective on student test activity*. Ph.D. dissertation, The University of Auckland.
- Cheng, L. (1997). How Does Washback Influence Teaching? Implications for Hong Kong, *Language and Education*, 11(1), 38–54. <https://doi.org/10.1080/09500789708666717>.
- Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24(3), 279–301.
- Cheng, L. (1999). Changing assessment: Washback on teacher perspectives and actions. *Teaching and Teacher Education*, 15(3), 253–271.
- Cheng, L., & Deluca, C. (2011). Voices from test-takers: Further evidence for language assessment validation and use. *Educational Assessment*, 16(2), 104–122.
- Cheng, L., Andrews, A., & Yu, Y. (2011). Impact and consequences of school-based assessment (SBA): Students' and parents' views of SBA in Hong Kong. *Language Testing*, 28(2), 221–249.
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching*, 48(4), 436–470.
- Chik, A., & Besser, S. (2011). International language test taking among young learners: A Hong Kong case study. *Language Assessment Quarterly*, 8(1), 73–91.
- Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education*, 27(2), 199–214.
- Davison, C., & Hamp-Lyons, L. (2009). The Hong Kong Certificate of Education: School-based assessment reform in Hong Kong English language education. In L.-Y. Cheng & A. Curtis (Eds.), *English language assessment and the Chinese learner*. Routledge.
- Dong, L. (2023). A study of IELTS's affective washback on Chinese students' learning goal, motivation, and anxiety. *Language Teaching Research Quarterly*, 9, 1–22. <https://doi.org/10.32038/ltrq.2018.09.01>
- Education Bureau. (2022). *Student Enrolment Statistics, 2021/22 (kindergartens, primary and secondary schools)*. Retrieved July 18, 2023 from https://www.edb.gov.hk/attachment/en/about-edb/publications-stat/figures/Enrol_2021.pdf
- Fields, A. (2013). *Discovering statistics using IBM SPSS statistics*.
- Fox, J., & Cheng, L. (2007). Did we take the same test? Differing accounts of the Ontario Secondary School Literacy Test by first and second language test-takers. *Assessment in Education: Principles, Policy and Practice*, 14(1), 9–26.
- Gee, J. P. (2007) *What video games have to teach us about learning and literacy*. Revised and updated edition. New York: Palgrave Macmillan.
- Green, A. (2006). Watching for washback: Observing the influence of the International English Language Testing System academic writing test in the classroom. *Language Assessment Quarterly*, 3, 333–368.
- Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education*, 14, 75–97.
- Gu, Y. H., & Lyu, Z. H. (2023). Backwash effect of English test for international communication on college English teaching. *Scholars International Journal of Linguistics and Literature*, 6(2), 136–140.
- Hong Kong Examinations and Assessment Authority (2019). Hong Kong Diploma of Secondary Education English Language Paper 4 guideline for candidates, Retrieved July 18, 2023 from https://www.hkeaa.edu.hk/DocLibrary/HKDSE/Subject_Information/eng_lang/EngDSESpeakingGuidelines2019.pdf
- Hong Kong Examinations and Assessment Authority. (2020). Hong Kong Diploma of Secondary Education Examination 2020 Registration for Category A and Category B subjects Instructions to Applicants (School Candidates). Retrieved July 18, 2023 from https://www.hkeaa.edu.hk/DocLibrary/HKDSE/Exam_Registration/Instructions_to_Applicants_SepReg_2020_SchoolCand_Eng.pdf
- Jacobs, J. E., & Eccles, J. S. (2000). Parents, task values, and real-life achievement-related choices. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic Motivation* (pp. 405–439). Academic Press.
- Khan, A., Hassan, N., & Ali, A. (2023). Implementing formative assessment in Malaysia: Teachers' viewpoints. *Language Testing in Focus: An International Journal*, 7, 28–41.
- Lam, H. P. (1994). Methodology washback— An insider's view. In D. Nunan, R. Berry, & V. Berry (Eds.), *Bringing about change in language education: Proceedings of the International Language in Education Conference 1994*. (pp. 83–102). University of Hong Kong.

- Maxmudjonova, X. (2023). *The role of washback in language teaching and testing*. Interpretation and Researches, 1(1). Retrieved July 18, 2023 from <http://interpretationandresearches.uz/index.php/iar/article/view/281>
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 243–256.
- Pan, Y. C. (2009). A review of washback and its pedagogical implications. *Journal of Science, Foreign Languages*, 25, 257–263.
- Pan, Y. C. (2014). Learner washback variability in standardized exit tests. *TESL-EJ*, 18(2), 1–30.
- Qi, L. (2004). *The intended washback effect of the National Matriculation English Test in China: Intentions and reality*. Foreign Language Teaching and Research Press.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, 22(2), 142–173.
- Qian, D. D. (2008). English language assessment in Hong Kong: A survey of practices, developments and issues. *Language Testing*, 25(1), 85–110.
- Samaie, M., & Valizadeh, P. (2023). The gap between teaching, assessment, and learning: EFL teachers' perceptions and strategies. *Language Testing in Focus: An International Journal*, 7, 12–27. <https://eurokd.com/doi/10.32038/ltf.2023.07.02>
- Saville, N. (2010). Developing a model for investigating the impact of language assessment. *Research Notes*, 42, 2–8. Retrieved July 18, 2023 from <https://www.cambridgeenglish.org/Images/23160-research-notes-42.pdf>
- Shih, C. M. (2007). A new washback model of students' learning. *The Canadian Modern Language Review*, 64(1), 135–162.
- Shih, C. M. (2010). The washback of the General English Proficiency Test on university policies: A Taiwan case study. *Language Assessment Quarterly*, 7, 234–254.
- Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education*, 14(1), 117–130.
- Shum, P. L. C., Tse, C. S., Hamamura, T., & Wright, S. C. (2023). The Effects of large-scale social movements on language attitudes: Cantonese and Mandarin in Hong Kong. *Journal of Language and Social Psychology*, 42(3), 249–274.
- Stoneman, B. W. H. (2006). *The impact of an exit English test on Hong Kong undergraduates: A study investigating the effects of test status on students' test preparation behaviours*. Unpublished doctoral dissertation, Hong Kong Polytechnic University.
- Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, 24(3), 302–321. <https://doi.org/10.1017/S095834401200016X>
- Tsagari, D. (2007b) *Review of washback in language testing: How has been done? What more needs doing?*. Lancaster University.
- Tsagari, D. (2020). Language assessment literacy: Concepts, challenges and prospects. In S. Hidri (Ed.), *Perspectives on Language Assessment Literacy: Challenges for improved student learning* (pp. 13–33). Routledge.
- Tsagari, D. (2007a). *Investigating the washback effect of a high-stakes EFL exam in the Greek context: Participants' perceptions, material design and classroom applications*.: PhD Thesis, Lancaster University. UK.
- Tsang, C. L. H. (2017). *Examining washback on learning from a sociocultural perspective: The case of a graded approach to English language testing in Hong Kong [Master Thesis]*. University College London: British Council Hong Kong.
- Tsang, C. L., & Isaacs, T. (2022). Hong Kong secondary students' perspectives on selecting test difficulty level and learner washback: Effects of a graded approach to assessment. *Language Testing*, 39(2), 212–238. <https://doi.org/10.1177/02655322211050600>
- Tsiplakou, S., & Tsagari, D. (2023). Language ideologies and washback effects in a high-stakes Greek language examination. *Research Papers in Language Teaching and Learning*, 13(1), 64–92.
- Watanabe, Y. (1992). Washback effects of college entrance examinations on language learning strategies. *JACET Bulletin*, 23, 175–194.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? *Preliminary Findings from Classroom-Based Research*, *Language Testing*, 13, 318–333.
- Watanabe, Y. (2004). Methodology in washback studies. *Washback in language testing: Research contexts and methods*. (pp. 19–36)
- Xie, Q., & Andrews, S. (2012). Do test design and uses influence test preparation? Testing a model of washback with structural equation modelling. *Language Testing*, 30(1), 49–70.
- Xie, Q. (2013). Does Test Preparation Work? Implications for Score Validity, *Language Assessment Quarterly*, 10(2), 196–218. <https://doi.org/10.1080/15434303.2012.721423>
- Yu, Y. (2010). *The washback effects of school-based assessment on teaching and learning: A case study*. Available from ProQuest Dissertations & Theses Global.
- Zhan, Y., & Andrews, S. (2014). Washback effects from a high-stakes examination on out-of-class English learning: Insights from possible self theories. *Assessment in Education: Principles, Policy & Practice*, 21(1), 71–89. <https://doi.org/10.1080/0969594X.2012.757546>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.