

Title of Project:

Washback Effect of Classroom-Based
Assessment from Learners' Perspectives

Researcher:

Yiwei Qin
coralqin@outlook.com

Masters-Degree Institution
University of Macau

Current Affiliation
University of Ottawa

Research Supervisor:

Prof. Cecilia Guanfang Zhao
University of Macau

TIRF Research Topic Investigated:

Language Assessment



Yiwei Qin

Motivation for the Research

The majority of washback studies have focused on the effect of large-scale high-stakes tests on different aspects of teaching and learning. Compared with the large number of washback studies that focus on teachers and their teaching practice, research from learners' perspective is scarce (Cheng, 2008; Spratt, 2005; Wall, 2000; Watanabe, 2004). More importantly, there has been little investigation into the effect of classroom-based assessment (CBA) on instructional and learning practices (Muñoz & Álvarez, 2010; Yu, 2010).

However, CBAs have long played a critical role in pedagogical practices and learning processes. As the designers and users of CBAs, classroom teachers often choose to use certain types of CBAs more often than other types, with an intention to facilitate and enhance students' learning. Although there is evidence of some shared views of washback effects brought about by the use of such CBAs among teachers and students, a mismatch might still exist between teacher and student perceptions, particularly with regard to the specific functions and effectiveness of different types of CBAs for language learning (Qi, 2007).

How assessment is integrated into classroom instruction and the process of language teaching and learning has been widely researched (e.g., Bachman & Palmer, 2010; Brindley, 1998; Davison, 2004; Davison & Leung, 2009; Llosa, 2005, 2007, 2008, 2012; McNamara, 2001; Rea-Dickins & Gardner, 2000; Rea-Dickins, 2001, 2004), but studies investigating how students perceive the intended washback effect of CBAs are scarce. Therefore, the purpose of this study is to investigate the washback effect of classroom-based assessments from the learners' perspective.

Research Questions

This study addressed the following research questions:



1. What are the most prevalent types of CBA adopted by English instructors in an English Language Centre (ELC) at a major English-medium university in an EFL context?
2. To what extent, if at all, do teacher perceptions of the relevant effectiveness of the CBAs align with student perceptions?
3. How does the use of CBAs influence students' second language (L2) learning in an intermediate-level EFL classroom over the course of one semester?

Research Methodology

This research was situated in the English Language Centre (ELC) at a major university in the Greater China region. ELC is a teaching unit that works with other departments to encourage and help students to improve their English skills, including general, academic, and professional English. All undergraduate students are distributed into different levels of compulsory English courses at ELC. Instead of one or two achievement tests (i.e., mid-terms, finals), ELC English courses evaluate students' English abilities based on their performance in various CBAs during the semester, including essays, presentations, discussions, projects, reading and listening quizzes, and textbook-based activities.

A mixed-methods design was adopted in this study. Data were collected through surveys, classroom observations and student interviews at various points over a semester. Specifically, teacher and student questionnaires were designed and administered to obtain an overview of the perceptions from both teachers and students about the nature, types, and functions of CBAs they work with in an EFL language center. Weekly classroom observations and student interviews over a semester supplemented the survey data and allowed the researcher to explore how exactly classroom-based assessments influenced L2 learners' language learning practices and attitudes.

Survey participants comprised of 16 teachers and 245 students. In addition, classroom observations (27.5 hours) of a class comprising 19 students were conducted. While four students from the class initially agreed to take part in the interview, only two of them completed the interviews over the course of ten weeks (14 interviews).

Three data collection instruments were developed for this study: a questionnaire, a classroom observation form, and a semi-structured interview protocol. A teacher questionnaire and a student questionnaire were used to elicit teacher and student perceptions of the frequency of various types of CBA used in the classroom setting and their perceived effectiveness for language learning. The items were designed based on Purpura and Turner's (2014) conceptualization of 17 types of CBAs and Alderson and Wall's (1993) conceptualization of washback effect in four main dimensions (i.e., an effect on students' overall learning, learning rate and sequence, degree or depth of learning, and attitude toward learning). A classroom observation form was developed to collect data on activities related to the CBA adopted by the teacher and the students' reactions to and performance on these CBAs over time. A semi-structured interview protocol was used to guide the weekly interviews with students that focused on three areas: students' understanding of the nature and function of specific CBAs; their attitudes toward the use of such CBAs and toward language learning; and their self-reported impact of various CBAs on their language learning (Cheng, 1998a; Qi, 2004; Yu, 2010).

Descriptive statistics and independent t-test were conducted to reveal teacher and student perceptions of the frequency of CBA use and its effects on students' learning, as well as the alignment (or lack thereof) between teachers' intended effect and students' perceived effect. The frequency of CBA use and the time duration of each type of CBA as noted down in the classroom observations were also analyzed. Content analysis of classroom observation data also focused on students' reactions to and performance on the CBAs implemented in the class. The observation data were used to triangulate data collected from the surveys and interviews. The analysis of the interview data consisted of three main stages. Firstly, the 14 interviews with students were transcribed and carefully reviewed to ensure

accuracy. Secondly, notable themes and patterns were initially identified through semantic coding (Dörnyei, 2007) of each student's interview data. Finally, a cross-case analysis was conducted, and three salient themes consistent across the narratives of the two students were identified.

Summary of Findings

Results from the analysis of both the survey and the classroom observation data showed that the most prevalent CBAs employed by classroom teachers in this particular EFL context were *oral questioning*, *class discussions*, and *projects*. Compared with the common assessment types (i.e., unit tests, midterms, and finals) the students experienced in their secondary education, the various CBAs used in this EFL course at the tertiary level clearly emphasized the learning process more than the score on one-time performance.

Comparisons of teacher and student perceptions of the washback effect of the three salient types of CBA showed that overall, there was a considerable amount of overlap between the teachers' intended washback effect on learning and the students' perceptions of that effect. In particular, *projects* were perceived by both parties as the most effective CBA that influenced students' overall English learning and their learning attitude. Despite the overall consistency between the teacher intentions and student perceptions, divergence was found with regard to specific areas of washback effect brought about by different CBAs. For instance, the teacher and student perceptions of the effectiveness of *class discussions* did not seem to match well. The analysis of the interview data indicated that the discrepancy between the teacher and student perceptions might be associated with the students' unfamiliarity with the CBA format and a general lack of adequate English-speaking ability. In addition, the students' personality also seemed to have played a role in their perceptions of the effect of certain CBA types. The role of personality may also be mediated by the students' self-regulation of their affective responses to the use of certain CBAs. To further complicate the picture, the students' understandings of the nature and functions of CBAs, and the weighing of specific CBAs in their final grades, might also be important factors influencing the students' learning attitudes and behavior. Last but not least, adequate learning opportunities provided by assessment users (i.e., classroom teachers) could also be an important factor in influencing the students' learning attitude and behavior.

Implications

Overall, the results from the current study carry a few important pedagogical implications. First, students' clear understanding of the nature and teachers' intended purposes of the CBAs are essential, if ever classroom teachers wanted to introduce and use CBAs to enhance teaching and learning quality. When implementing CBAs, classroom teachers should, therefore, explain explicitly and thoroughly the intended purposes of using one CBA over another. Only when students fully understand how they can benefit from their work on these CBAs, do they become active participants in the learning process, and assume more responsibilities as independent and autonomous learners. Second, when designing CBAs, teachers may want to provide more learning opportunities by extending the time needed for each task, offering more constructive feedback, and providing collaborative working opportunities in classrooms. Third, CBAs that constitute a higher percentage toward the final grades may provide stronger motivation for students' learning. This does not mean that teachers should simply assign a higher percentage of the final score to an ideal CBA, but rather that students should be involved in the design and evaluation process. For example, with teacher assistance, students could collectively construct a rubric that expresses performance expectations and evaluative criteria in language meaningful to students themselves. Overall, teachers should use various techniques to develop a "learning community" in the classroom and to encourage students to "become more actively interested in the process of learning itself" (Byon 2005, p. 186).



References

- Airasian, P. W. (2005). *Classroom assessment: Concepts and applications* (5th ed.). McGraw-Hill.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280–297. <https://doi.org/10.1177/026553229601300304>
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics*, 14(2), 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Andrews, S., & Fullilove, J. (1994). Assessing spoken English in public examinations—Why and how? In J. Boyle, & P. Falvey (Eds.) *English language testing in Hong Kong* (pp. 57–85). The Chinese University Press.
- 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)
- Angrosino, M. (2007). *Doing ethnographic and observational research*. SAGE Publications Ltd. <https://dx.doi.org/10.4135/9781849208932>
- Assessment Reform Group. (1999). *Assessment for learning: Beyond the black box*. Nuffield Foundation. [http:// https://doi.org/10.13140/2.1.2840.1444](http://https://doi.org/10.13140/2.1.2840.1444)
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- Bailey, K. M. (1999). *Washback in language testing*. Educational Testing Service.
- Baksh, A., Mohd Sallehudin, A., Tayeb, Y., & Norhaslinda, H. (2016). Washback effect of school-based English language assessment: A case-study on students' perceptions. *Pertanika Journal of Social Sciences & Humanities*, 24(3), 1069–1086.
- Black, P., Harrison, C., & Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. McGraw-Hill.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, policy & practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Brenner, M., Brown, J., & Canter, D. V. (Eds.). (1985). *The Research interview: Uses and approaches*. Academic Press.
- Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing*, 15(1), 45–85. <https://doi.org/10.1177/026553229801500103>
- Brown, J. D., & Hudson, T. (2012). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653–675. <https://doi.org/10.2307/3587999>



- Bryman, A. (1992). Quantitative and qualitative research: Further reflections on their integration. In Brannen, J. (Ed.), *Mixing methods: Qualitative and quantitative research* (pp. 57–80). Routledge. <https://doi.org/10.4324/9781315248813>
- Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in education and teaching international*, 44(1), 57–66. <https://doi.org/10.1080/14703290601081332>
- Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education*, 69(6), 963–976. <https://doi.org/10.1007/s10734-014-9816-z>
- Center for Educational Research and Innovation. (2005). *Formative assessment: Improving learning in secondary classrooms*. Organisation for Economic Co-operation and Development. <https://doi.org/10.1787/9789264007413-en>
- Chapman, D. W., & Snyder Jr, C. W. (2000). Can high stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development*, 20(6), 457-474. [https://doi.org/10.1016/S0738-0593\(00\)00020-1](https://doi.org/10.1016/S0738-0593(00)00020-1)
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and education*, 11(1), 38–54. <https://doi.org/10.1080/09500789708666717>
- Cheng, L. (1998a). *The washback effect of public examination change on classroom teaching: An impact study of the 1996 Hong Kong Certificate of Education in English on the classroom teaching of English in Hong Kong secondary schools* [Unpublished doctoral dissertation]. University of Hong Kong.
- Cheng, L. (1998b). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24(3), 279–301. [https://doi.org/10.1016/S0191-491X\(98\)00018-2](https://doi.org/10.1016/S0191-491X(98)00018-2)
- Cheng, L. (2001). Washback studies: Methodological considerations. *Curriculum Forum* 10(2), 17–32.
- Cheng, L. (2003). Looking at the impact of a public examination change on secondary classroom teaching: A Hong Kong case study. *The Journal of Classroom Interaction*, 38(1), 1–10.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.
- Cheng L. (2008) Washback, impact and consequences. In N. H. Hornberger (Ed.), *Encyclopedia of Language and Education* (2nd ed., pp. 2479–2494). Springer.
- 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. <https://doi.org/10.1017/S0261444815000233>
- Cheng, L., & Watanabe, Y. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Routledge.



- Cheng, L., Andrews, S., & 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. <https://doi.org/10.1177/0265532210384253>
- Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. Jossey-Bass.
- Clapham, C. (2000). Assessment and testing. *Annual Review of Applied Linguistics*, 20, 147–161. <https://doi.org/10.1017/S0267190500200093>
- Creswell, J. W., Tashakkori, A., Jensen, K. D., & Shapley, K. L. (2003). Teaching mixed methods research: Practices, dilemmas, and challenges. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 91–110). SAGE.
- Dann, R. (2012). *Promoting assessment as learning: Improving the learning process*. Routledge.
- Dann, R. (2014). Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in education: Principles, policy & practice*, 21(2), 149–166. <https://doi.org/10.1080/0969594X.2014.898128>
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305–334. <https://doi.org/10.1191/0265532204lt286oa>
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43(3), 393–415. <https://doi.org/10.1002/j.1545-7249.2009.tb00242.x>
- Darlington, Y., & Scott, D (2002). *Qualitative research in practice: Stories from the field*. Crowsnest.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). Corwin.
- Erickson, F. (1996). On the evolution of qualitative approaches in educational research: from Adam's task to Eve's. *Australian Educational Researcher*, 23(2), 1–15. <https://doi.org/10.1007/BF03219617>
- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng, & Y. Watanabe (Eds.) *Washback in language testing: Research contexts and methods* (pp. 191–210). Routledge.
- Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational researcher*, 18(9), 27–32. <https://doi.org/10.3102/0013189X018009027>
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction* (6th ed.). Longman.



- Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-session language courses. *Assessment in Education: Principles, Policy & Practice*, 14(1), 75–97. <https://doi.org/10.1080/09695940701272880>
- Hammersley, M., & Atkinson, P. (1983). *Ethnography: Principles in practice*. Tavistock.
- Hamp-Lyons, L. (1998). Ethical test preparation practice: The case of the TOEFL. *TESOL Quarterly*, 32(2), 329–337. <https://doi.org/10.2307/3587587>
- Hanrahan, S. J., & Isaacs, G. (2001). Assessing self-and peer-assessment: The students' views. *Higher Education Research & Development*, 20(1), 53–70. <https://doi.org/10.1080/07294360123776>
- Hawkey, R. A. (2006). *Impact theory and practice: Studies of the IELTS test and Progetto Lingue 2000*. Cambridge University Press.
- Hernandez, R. (2012). Collaborative learning: Increasing students' engagement outside the classroom. *US-China Education Review A* 2(9), 804–812. <https://doi.org/10.17265/2161-623X/2012.09A.006>
- Huerta-Macías, A. (2007). Alternative assessment: Responses to commonly asked questions. In J. Richards, & W. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 338–343). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667190.048>
- Hughes, A. (1993). *Backwash and TOEFL 2000*. [Unpublished manuscript]. University of Reading.
- Hughes, I. E., & Large, B. J. (1993). Staff and peer-group assessment of oral communication skills. *Studies in Higher Education*, 18(3), 379–385. <https://doi.org/10.1080/03075079312331382281>
- Iredale, C. (1990). Pupil's attitudes towards GASP (Graded Assessments in Science Project). *School science review*, 72(258), 133–137.
- Jorgensen, D. L. (1989). *Participant Observation*. SAGE. <https://dx.doi.org/10.4135/9781412985376>
- Lam, H. P. (1993). *Washback: Can it be quantified? A study on the impact of English examinations in Hong Kong* [Unpublished master's dissertation]. University of Leeds.
- Lam, R. (2016). Assessment as learning: Examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education*, 41(11), 1900–1917. <https://doi.org/10.1080/03075079.2014.999317>
- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, 9(2), 202–220.
- Llosa, L. (2005). *Building and supporting a validity argument for a standards-based classroom assessment of English proficiency* (Publication No. 3181711) [Doctoral dissertation, University of California]. ProQuest Dissertations and Theses Global.



- Llosa, L. (2007). Validating a standards-based classroom assessment of English proficiency: A multitrait-multimethod approach. *Language testing*, 24(4), 489–515.
<https://doi.org/10.1177/0265532207080770>
- Llosa, L. (2008). Building and supporting a validity argument for a standards-based classroom assessment of English proficiency based on teacher judgments. *Educational Measurement: Issues and Practice*, 27(3), 32–42. <https://doi.org/10.1111/j.1745-3992.2008.00126.x>
- Llosa, L. (2012). Assessing English learners' progress: Longitudinal invariance of a standards-based classroom assessment of English proficiency. *Language Assessment Quarterly*, 9(4), 331–347.
<https://doi.org/10.1080/15434303.2012.721422>
- Lumley, T., & Stoneman, B. (2000). Conflicting perspectives on the role of test preparation in relation to learning? *Hong Kong Journal of Applied Linguistics*, 5(1), 50–80.
- Madaus, G. F. (1988). The distortion of teaching and testing: High-stakes testing and instruction. *Peabody Journal of Education*, 65(3), 29–46.
<https://doi.org/10.1080/01619568809538611>
- Mckay, S. L. (2006). *Researching second language classrooms*. Routledge.
<https://doi.org/10.4324/9781410617378>
- McNamara, T. (1996). *Measuring second language performance*. Longman.
- McNamara, T. (2000). *Language testing*. Oxford University Press.
- McNamara, T. (2001). Rethinking alternative assessment. *Language testing*, 18(4), 329–332.
<https://doi.org/10.1177/026553220101800401>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp 13–103). Macmillan.
- Messick, S. (1996). Validity and washback in language testing. *Language testing*, 13(3), 241–256.
<https://doi.org/10.1177/026553229601300302>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). SAGE Publications Ltd.
- Morrow, K. (1986). The evaluation of tests of communicative performance. In M. Portal (Ed.), *Innovations in language testing: Proceedings of the IUS/NFER conference* (pp. 1–13). NFER-Nelson.
- Muñoz, A. P., & Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language testing*, 27(1), 33–49. <https://doi.org/10.1177/0265532209347148>
- Noble, A. J., & Smith, M. L. (1994). *Old and New Beliefs about Measurement-Driven Reform: "The More Things Change, the More They Stay the Same"* (ED378228). ERIC.
<https://eric.ed.gov/?id=ED378228>



- Patton, M. Q. (1987). *How to use qualitative methods in evaluation*. SAGE.
- Pearson, I. (1988). Tests as levers for change. In D. Chamberlain, & R. J. Baumgardner (Eds.), *ESP in the classroom: Practice and evaluation* (pp. 98–107). Modern English Publications.
- Perrone, M. (2011). The effect of classroom-based assessment and language processing on the second language acquisition of EFL students. *Journal of Adult Education*, 40(1), 20–33.
- Popham, W. J. (1987). The merits of measurement-driven instruction. *The Phi Delta Kappan*, 68(9), 679–682.
- Purpura, J. E. (2004). *Assessing grammar*. Cambridge University Press.
- Purpura, J. E. (2009). The impact of large-scale and classroom-based language assessment on the individual. In L. Taylor & C. J. Weir (Eds.), *Language testing matters: Investigating the wider social and educational impact of assessment* (pp. 301–325). Cambridge University Press.
- Purpura, J. E., & Turner, C. E. (2014, October 12–14). *A learning-oriented assessment approach to understanding the complexities of classroom-based language assessment* [Paper presentation]. Roundtable on learning-Oriented Assessment in Language Classrooms and Large-Scale Assessment Contexts. Teachers College, Columbia University, New York, United States.
- Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng & Y. Watanabe (Eds.) *Washback in language testing: Research contexts and methods* (pp. 171–190). Routledge.
- 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. <https://doi.org/10.1080/09695940701272856>
- Rea-Dickins, P. (2000). Classroom assessment. In T. Hedge (Ed.), *Teaching and learning in the language classroom* (pp. 375-401). Oxford University Press.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language testing*, 18(4), 429–462. <https://doi.org/10.1177/026553220101800407>
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249–258. <https://doi.org/10.1191/0265532204lt283ed>
- Rea-Dickins, P. (2007). Classroom-based assessment: Possibilities and pitfalls. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 505–520). Springer. https://doi.org/10.1007/978-0-387-46301-8_36
- Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 17(2), 215–243. <https://doi.org/10.1177/026553220001700206>
- Shih, C. M. (2007). A new washback model of students' learning. *Canadian Modern Language Review*, 64(1), 135–161. <https://doi.org/10.3138/cmlr.64.1.135>



- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal*, 76(4), 513–521. <https://doi.org/10.2307/330053>
- Shohamy, E. (1993). *The power of tests: The impact of language tests on teaching and learning* (ED362040). ERIC. <https://files.eric.ed.gov/fulltext/ED362040.pdf>
- Sluijsmans, D. M. A. (2002). *Student involvement in assessment: The training of peer assessment skills* [Unpublished doctoral dissertation]. Open University of the Netherlands.
- Smith, M. L. (1991). Meanings of test preparation. *American Educational Research Journal*, 28(3), 521–542. <https://doi.org/10.3102/00028312028003521>
- Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL quarterly*, 42(2), 181–207. <https://doi.org/10.1002/j.1545-7249.2008.tb00115.x>
- 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. <https://doi.org/10.1191/1362168805lr152oa>
- 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* [Doctoral dissertation, The Hong Kong Polytechnic University]. Pao Yue-kong Library. <http://hdl.handle.net/10397/2720>
- Taylor, S. J., & Bogdan, R. (1998). *Introduction to qualitative research methods: A guidebook and resource* (3rd ed.). Wiley.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281–294. <https://doi.org/10.1080/09695940701591867>
- Tsagari D., & Cheng L. (2017). Washback, impact, and consequences revisited. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment* (3rd ed., pp. 359–372). Springer.
- Turner, E. C. (2012). Classroom assessment. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 65–78). Routledge.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Ablex.
- Wall, D. (1996). Introducing new tests into traditional systems: Insights from general education and from innovation theory. *Language Testing*, 13(3), 334–354. <https://doi.org/10.1177/026553229601300307>
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System*, 28(4), 499–509. [https://doi.org/10.1016/S0346-251X\(00\)00035-X](https://doi.org/10.1016/S0346-251X(00)00035-X)



- Wall, D. (2005). *The impact of high-stakes examinations on classroom teaching: A case study using insights from testing and innovation theory*. Cambridge University Press.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, 13(3), 318–333.
<https://doi.org/10.1177/026553229601300306>
- Watanabe, Y. (1997). *The washback effects of the Japanese university entrance examinations of English: Classroom-based research* (uk.bl.ethos.389858) [Doctoral dissertation, Lancaster University]. EThOS. <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.389858>
- Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng & Y. Watanabe (Eds.) *Washback in language testing: Research contexts and methods* (pp. 41–58). Routledge.
- Xu, Q., & Liu, J. (2018). *A study on the washback effects of the test for English majors (TEM): Implications for testing and teaching reforms*. Springer.
- Yu, Y. (2010). *The washback effects of school-based assessment on teaching and learning: A case study* [Doctoral dissertation, University of Hong Kong]. The HKU Scholars Hub.
<http://hdl.handle.net/10722/65273>