

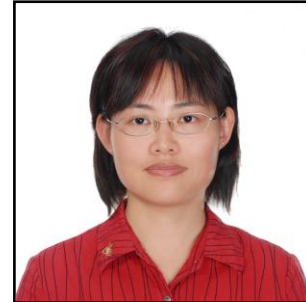


**Title of Project:**

Washback Effects of the Reformed CET-4 on College English Teaching and Learning in China: Students' Perspectives

**Researcher:**

Zhiling Wu  
Indiana University of Pennsylvania  
[fkkm@iup.edu](mailto:fkkm@iup.edu)



Zhiling Wu

**Research Supervisor:**

Dr. Michael M. Williamson

---

**Project Summary**

The title of my dissertation is *Washback Effects of the Reformed CET-4 on College English Teaching and Learning in China: Students' Perspectives*. Washback is defined as the influence of testing on teaching and learning, according to Bailey (1996). Given that there are not many empirical studies on washback, Alderson and Wall (1993) made an appeal for more studies on washback and developed 15 washback hypotheses to direct future research. An increasing number of empirical studies appeared after the call by Alderson and Wall. However, the majority of these studies focuses primarily on teacher's perspectives and how English teaching has been influenced by high-stakes tests, (e.g., Cheng, 2005; Manjarres, 2005; Qi, 2004; Shohamy, 2001; Watanabe, 2004). However, the perspectives of the test takers, who are the most immediate stakeholders, have been largely under-researched. Equally under-researched are the washback effects of language tests at the college level because most reported research has been conducted on tests before the tertiary level (e.g., Andrews, Fullilove & Wong, 2002; Huhta, Kalaja, & Pitkanen-Huhta, 2006; Salehi, Mustapha & Yunus, n.d.; Watanabe, 1996). What is more important theoretically, according to unified validity theory (Messick, 1996), is evidence of washback that is related to the consequential aspect of construct validation enquiry.

Hence, the present study, from students' perspectives, examines how the reformed national College English Test Band 4 (CET-4) affects English teaching and learning at the college level in China. The CET-4 is the only national test for non-English majors at the college level. It is administered twice a year: June and December. With its inception in 1987, it has drawn the largest number of test takers in the world (Jin and Yang, 2006). In 2006, for instance, more than 13 million students took the CET-4 (Zheng and Cheng, 2008). In response to "the pressing social need for college and university graduates with a strong communicative competence in English" (Jin and Yang, p. 21) and the cry to improve the washback effects of the CET-4, the Ministry of Education in China and the College English Testing Committee Commission reformed the CET-4 and implemented the reformed version nationwide in January 2007.

This study was conducted in three universities in Shanghai, the biggest city in China. Though



selected on the basis of convenience sampling, they represent three levels of prestige in the Chinese hierarchy of higher education – University A, highly reputed; University B, intermediate; and University C, ordinary. A four-phase mixed method is employed. In Phase I, a survey regarding the influence of the reformed CET-4 on students' college English teaching and learning was administered in three different classes at each university in May 2010. In Phase II, from among the students who had indicated their interest in participating in the follow-up interview at the end of the survey, ten students from each university (eight planned to take the CET-4 in June and two in December) were selected based on gender, major, English classes registered, and English exam grades at matriculation.

One problem for Phase II is that the normal time for students at University A to take the CET-4 was December, whereas it was June for the other two and most of the universities in China. Thanks to a doctoral dissertation grant (DDG) from The International Research Foundation (TIRF) the study was then expanded to Phase III and IV. In Phase III, a slightly modified survey was administered in November 2010 to two classes who were going to take the CET-4 in December in University A. In Phase IV, four volunteers in each class were selected to participate in the follow-up interview. What is more, the six students who were interviewed in Phase II, yet were going to take the CET-4 in December at this time were contacted to complete the survey and were asked to be interviewed again. All 14 students in Phase IV were also asked to keep journals or self-recordings about their CET-4 test-taking experience. Therefore, altogether, 414 participants were surveyed (5 were surveyed twice), 34 were interviewed (6 were interviewed twice), and 7 submitted their journals or self-recordings.

There are three major research questions in this study:

1. What are college students' beliefs about English learning; what are their expectations and experiences of studying English at college, and what are their perceptions of the reformed CET-4?
2. What plans did the students make to prepare for taking the CET-4, and how did they implement their plans?
3. To what extent has the reformed CET-4 influenced students' perceptions of the practices of English teaching and learning at the college level?

My major findings are discussed according to each research question and beginning with Question 1. Despite the fact that almost all the college students in the study acknowledged the essential status of English in the development of the global economy, politics, cultural exchange and education, English is largely considered a means to a utilitarian end such as college graduation, employment, and/or potential assistantship/scholarship. In general, they value English, but they do not value the College English course. Moreover, surprisingly, a number of participants knew little about the modified policy and up-to-date information of the reformed CET-4 even though they were going to take the test within a month. Then, there were four major reasons why 65.5% of the students in the three universities spent only 0.5 hour after class each day on English and paid less attention to English studying at the college level. They were 1) college students' heavy study load, 2) lack of teacher guidance, 3) the low status of the College



English course, and 4) perceived relatively low pressure caused by the reformed CET-4.

For research Question 2, almost half of the interviewees bought a CET-4 vocabulary book of some kind and/or copies of the past CET-4 test or simulated tests to prepare for the exam. Approximately 84% of the 414 participants claimed that they planned to prepare for the CET-4. The majority of the students planned to memorize the CET-4 vocabulary and take the simulated tests or tests from previous years. According to the self-recording and journal participants, their preparation focus was mainly on CET-4 vocabulary, listening and banked cloze. Overall, it is still safe to conclude that the reformed CET-4 has had much washback effect on college English teaching and learning.

For research Question 3, this study found that, in particular, three positive washback effects targeted by the reform of the CET-4 have been achieved: 1) greater use of English as the language of instruction, 2) more use of computers to assist teachers in and out of the classroom, and 3) greater emphasis on English listening skills. Besides, the data suggest that the reformed CET-4 has changed what teachers teach and what students learn in and out of English classes. Inferential statistics show that students at University C did significantly more fast reading, mock/past tests (in class and out of class), and intensive reading than students at Universities A and B. Among students who registered to take the CET-4 in June, teachers at University C mentioned the CET-4 significantly more often in class than their counterparts among the three universities. At University B, test takers who registered for the CET-4 in December did significantly more listening practice in class than their counterparts at Universities A and C. Teachers at University B, at this time mentioned the test most frequently in class, too. Teachers at University A, in contrast, gave students significantly more writing practice and less fast reading or intensive reading than their counterparts at Universities B and C. University A also gave students more demanding listening materials than the other two universities. So, in accordance with the findings by Alderson and Hamp-Lyons (1996), the reformed CET-4 has different amounts and types of washback on some teachers and learners than on other teachers and learners. What is more, this study shows that the lower the ranking of the university, the stronger the washback effects are.

Further analyses, however, indicate that the reformed CET-4 has not changed how teachers teach or how students learn. The requirement of using assigned textbooks by the Ministry of Education seems to be the major barrier. No substantial two-way interaction has taken place in English classes. A number of participants did not even know the existence of the CET- Spoken English Test (SET). Writing is, in fact, another area interviewees reported that they did not receive sufficient instruction. In other words, the reformed CET-4 did not have the desired washback effects on English teaching and writing as that claimed in the survey data. Except University A, both teachers and students at Universities B and C skipped writing instruction and exercises. Memorizing formulaic writing templates was usually how students prepared for writing tasks.

In order to improve the washback effects of the reformed CET-4 and students' English communicative competence, the researcher proposes that measures (e.g., CET-4 workshops) should be taken to guarantee that students are acquainted with the policies of the informed CET-



4. Moreover, CET-SET should receive more publicity. It should also be integrated into the CET-4 and made accessible to all college students. For this purpose, oral English classes of different levels should be offered as a separate class in Chinese universities. For writing, considering the burden teachers may have with large classes, writing and grading assistance would be beneficial for both English teachers and students. Funding permitting, teaching assistants could be assigned to teachers to help them grade essays. Writing centers should be established to offer students extra help with their English writing skills.

Last but not least, given the uneven development of English fundamentals acquired prior to college, the variety of motivation for studying English in college, and different English proficiency levels demanded by different majors, it would be wise to offer more diverse English courses other than the uniformed, one-size-fits-all college English course. The English courses offered at University A could be a good model to follow. A more challenging version of CET-4 can be administered to students at top universities because they have teachers and students with comparatively stronger backgrounds, in addition to richer teaching and learning resources. Of course, none of the positive washback effects that were intended to achieve can take place easily if support services, such as self-access English learning labs, improved English education prior to college, and intensive teacher training programs, are not provided.

To conclude, this study reinforces that students' perspectives should be taken into account in washback studies. Validating a test and trying to improve language education through a test's positive washback is a long-term and on-going task. It cannot be executed successfully without a collective effort that involves of all the stakeholders.



## References

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*(3), 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129.
- Alderson, J. C. (2004). Foreword. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. ix-xii). Mahwah, NJ: Lawrence Erlbaum.
- American Psychological Association (1985). *Standards for educational and psychological testing*. Washington, DC: Author.
- 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.
- Bachman, F. L. (1990). *Fundamental considerations in language testing*. Oxford, UK: Oxford University Press.
- Bachman F. L. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 41-72). Ottawa, Canada: University of Ottawa Press.
- Bachman, F. L., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, UK: Oxford University Press.
- Bailey, K. M. (1990). The use of diary studies in teacher education programs. In J.C Richards, & D. Nunan (Eds.), *Second language teacher education* (pp. 215-226). Cambridge, UK: Cambridge University Press.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept. *Language Testing, 13*(3), 257-279.
- Biggs, J. B. (1995). Assumptions underlying new approaches to educational assessment. *Curriculum Forum, 4*(2), 1-22.
- Biggs, J. B. (Ed.) (1996). *Testing: To educate or to select? Education in Hong Kong at the cross-roads*. Hong Kong: Hong Kong Educational.



- Bourdieu, P., & Thompson, J. B. (1991). *Language and symbolic power*. Cambridge, MA: Harvard University Press.
- Bracey, G. W. (1987). Measurement-driven instruction: Catchy phrase, dangerous practice. *Phi Delta Kappa*, 68, 683-686.
- Brannen, J. (1992). Combining qualitative and quantitative approaches: An overview. In J. Brannen (Ed.), *Mixing methods: Qualitative and quantitative research* (pp.3-37). Burlington, VT: Ashgate.
- Bray, M., & Steward, L. (1998). *Examination systems in small states: Comparative perspectives on policies, models and operations*. London, UK: Commonwealth Secretariat.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *Japan Association for Language Teaching Journal*, 10(1), 12-42.
- Burrows, C. (1998). *Searching for washback: An investigation of the impact on teachers of the implementation into the adult migrant English program of the assessment of the certificates in spoken and written English* (Unpublished doctoral dissertation). Macquarie University, Sydney.
- Cai, J. (2006). *ELT at tertiary level in China: Review, reflection and research*. Fudan, Shanghai: Fudan University Press.
- Canale, M. (1983). On some dimensions of language proficiency. In J.W. Oller, Jr. (Ed.), *Issues in language testing research* (pp. 333-342). Rowley, MA: Newbury House.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approach to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Carroll, J. B. (1968). The psychology of language testing. In A. Davies (Ed.), *Language testing symposium: A psycholinguistic approach* (pp. 46-69). London, UK: Oxford University Press.
- Celce-Murcia, M. (2001). *Teaching English as a second or foreign language*. (3rd ed.). Boston, MA: Heinle & Heinle.
- Chapman D. W., & Snyder, C. W. Jr. (2000). Can high stakes national testing improve instruction? Reexamining conventional wisdom. *International Journal of Educational Development*, 20(6), 457-474.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge, UK: Cambridge University Press.



- Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, 25(1), 15-37.
- Cheng, L. (2009). The history of examinations: Why, how, what and whom to select? In L. Cheng, & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 13-25). New York, NY: Routledge.
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp.19–36). Mahwah, NJ: Lawrence Erlbaum.
- Cheng, L., & Curtis, A. (2009). The realities of English language assessment and the Chinese learner in China and beyond. In L. Cheng, & A. Curtis (Eds.), *English Language assessment and the Chinese learner* (pp. 3-12). New York, NY: Routledge.
- 全国约有 3 亿多人在学英语。 [China has approximately 30 million English learners]. (2006, April 18). Retrieved from <http://opinion.people.com.cn/GB/8213/62024/430907.html>
- 中国大学排行榜: 上海 [China 2010 National University Rankings: Shanghai]. (2010). Retrieved from [http://rank2010.netbig.com/citypaihang\\_9.html](http://rank2010.netbig.com/citypaihang_9.html)
- Corbett, H. D., & Wilson, B. L. (1991). *Testing, reform and rebellion*. Norwood, NJ: Ablex.
- Creswell, J.W. (2008). *Educational research: Planning, conducting, and evaluating qualitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M., & Hanson, W. (2003). Advanced Mixed Methods Research Designs. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 619-637). Thousand Oaks, CA: Sage.
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Crystal, D. (2003). *English as a global language*. (2nd ed.). Cambridge, UK: Cambridge University Press.
- Davies, A. (1977). The construction of language tests. In J. P. B. Allen, & A. Davies (Eds.), *Testing and experimental methods* (pp. 38-104). London, UK: Oxford University Press.
- Davies, A. (1990). *Principles of language testing*. Oxford, UK: Basil Blackwell.



- Davies, A. (2007). Assessing academic English language proficiency: 40 years of UK language tests. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 73-86). Ottawa, Canada: University of Ottawa Press.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Eckstein, M. A., & Noah, H. J. (Eds.). (1992). *Examinations: Comparative and international studies*. New York, NY: Pergamon Press.
- English, F.W. (1992). *Deciding what to teach and test: Developing, aligning, and auditing the curriculum*. Newbury Park, CA: Corwin Press.
- Foster, P. J. (1992). Commentary. In M.A. Eckstein, & H. J. Noah (Eds.), *Examinations: Comparative and international studies* (pp. 121–126). Oxford, UK: Pergamon Press.
- Frederickson, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychology*, 39(3), 193-202.
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). New York, NY: Teachers College Press.
- Gao, Y. (2003). Empirical study and open-door policy are needed for foreign language teaching reform. *Foreign Language Teaching and Research*, 99(3), 222-223.
- Gates, S. (1995). Exploiting washback from standardized tests. In J. D. Brown, & S. O. Yamashita (Eds.), *Language testing in Japan* (pp. 101-106). Tokyo, Japan: Japanese Association for Language Teaching.
- Green, A. B. (2007). *IELTS washback in context: Preparation for academic writing in higher education*. Cambridge, UK: Cambridge University Press.
- Gregory, S. (2012, January 4). Why doesn't Serena Williamson love tennis? Retrieved from <http://keepingscore.blogs.time.com/2012/01/04/why-doesnt-serena-williams-love-tennis/>
- Ha, P. L. (2008). *Teaching English as an international language: Identity, resistance and negotiation*. Clevedon, UK: Multilingual Matters.
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14(3), 295-303.
- Herman, J. L., & Golan, S. (1991). *Effects of standardized testing on teachers and learning: Another look* (CSE Technical Report No. 334). Los Angeles, CA: University of California, CRESST.





- Hu, C.T. (1984). The historical background: examinations and control in pre-modern China. *Comparative Education*, 20(1), 7-26.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge, UK: Cambridge University Press.
- Hughes, A. (1993). *Backwash and TOEFL 2000*. Unpublished manuscript, University of Reading, Reading, Berkshire, United Kingdom.
- Hughes, A (2003). *Testing for language teachers* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Huhta, A., Kalaja, P., & Pitkanen-Huhta, A. (2006). Discursive construction of a high-stakes test: The many faces of a test-taker. *Language Testing*, 23(3), 326-350.
- Jin, Y. (2006). Improving validity and washback: A washback study on the CET-4 and CET-6. *Foreign Language World*, (6), 65-73.
- Jin, Y. (2010). The national College English Testing Committee. In L. Cheng, & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 44-59). New York, NY: Routledge.
- Jin, Y., & Yang, H. (2006). The English proficiency of college and university students in China: As reflected in the CET. *Language, Culture and Curriculum*, 19(1), 21-36.
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London, UK: Longman.
- Lam, A. (2002). English in education in China: Policy changes and learners' experiences. *World Englishes*, 21(2), 245-256.
- Larsen-Freeman, D. (2000). *Techniques and principles in language teaching: Teaching techniques in English as a second language*. (2nd ed.). Oxford, UK: Oxford University Press.
- Lee, J. F., & VanPatten, B. (1995). *Making communicative language teaching happen*. New York, NY: McGraw-Hill.
- Li, Q., & Wang, X. (2003). Inspiration of basic English education reform for college English. *Foreign Language Teaching and Research*, (5), 382-383.
- Littlewood, W. (1981). *Communicative language teaching: An introduction*. Cambridge, UK: Cambridge University Press.



- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and criteria. *Educational Researcher*, 20(8), 15-21.
- Luxia, Q. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, 22(2), 142-173.
- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum: Eighty-seventh yearbook of the national society for the study of education* (pp. 83-121). Chicago, IL: University of Chicago Press.
- Manjarrés, N. B. (2005). Washback of the foreign language test of the state examinations in Colombia: A case study. *Arizona Working Papers in SLAT*, 12, 1-19.
- McNamara, T., & Roever, C. (2006). *Language Testing: The social dimension*. Oxford, UK: Blackwell.
- McKay, S. L. (2002). *Teaching English as an international language: Rethinking goals and approaches*. Oxford, UK: Oxford University Press.
- Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer, & H. I. Braun (Eds.), *Test validity* (pp.33-45). Hillsdale, NJ: Lawrence Erlbaum.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- 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). London, UK: NFER/Nelson.
- National College English Testing Committee (1985). *Syllabus for College English Test*. [Brochure]. Shanghai, China: Shanghai Language Education Press.
- National College English Testing Committee (1999). *Syllabus for College English Test*. [Brochure]. Shanghai, China: Shanghai Language Education Press.
- National College English Testing Committee (2006). *Syllabus for College English Test*. [Brochure]. Shanghai, China: Shanghai Language Education Press.
- National College English Testing Committee (2006). *College English Test - Spoken English Test (CET-SET) syllabus and sample test*. [Brochure]. Shanghai, China: Shanghai Language Education Press.



- National College English Testing Committee (2006). *CET-4 syllabus and sample test paper (revised version)*. [Brochure]. Shanghai, China: Shanghai Language Education Press.
- Nobel, A. J., & Smith, M. L. (1994). *Measurement-driven reform: Research on policy, practice, repercussion*, (CES Technical Report No. 381). Tempe, Arizona: Arizona State University, CSE.
- Oneill, P., Murphy, S., Huot, B., & Williamson, M. (2006). What high school teachers in three states say about different kinds of mandated state writing tests. *Journal of Writing Assessment*, 2(2), 81–108.
- Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 351-383). Thousand Oaks, CA: 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). London, UK: Modern English.
- Popham, W. J. (1983). Measurement as an instructional catalyst. In R. B. Ekstrom (Ed.), *Measurement, technology, and individuality in education: Proceedings of the 1982 ETS Invitational Conference: No.17. New directions for testing and measurement* (pp.19-30). San Francisco, CA: Jossey-Bass.
- Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- Popham, W. J. (1993). Measurement-driven instruction as a “quick-fix” reform strategy. *Measurement and Evaluation in Counseling and Development*, 26(1), 31-34.
- Potter, J. (2003). Discourse analysis and discursive psychology. In P. M. Camic, J. E. Rhodes, & L. Yardley, (Eds.), *Qualitative research in psychology: Expanding perspectives in methodology and design* (pp.73-94). Washington, DC: American Psychological Association.
- Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 171-190). Mahwah, NJ: Lawrence Erlbaum.
- Rea-Dickins, P. (1997). ‘So, why do we need relationships with stakeholders in language testing? A view from the UK’. *Language Testing*, 14(3), 304-314.
- Richards, J. C., Platt, J., & Platt, H. (1992) *Longman dictionary of language teaching and applied linguistics*. Harlow, Essex: Longman .



- Ross, J. S. (2008). Language testing in Asia: Evolution, innovation, and policy challenges. *Language Testing*, 25(1), 5-13.
- Salehi, H., Mustapha, R., & Yunus, M. M. (n.d.). The perceptions and attitudes of Iranian high school English teachers, students, and parents towards the entrance examination of universities (EEOU): Implications on English learning and teaching. *Prosiding Seminar Penyelidikan Siswazah*, 75-84.
- Savignon, S.J. (1997). *Communicative competence: Theory and classroom practice*. New York, NY: McGraw-Hill.
- Saville, N., & Hawkey, R. (2004). The IELTS impact study: Investigating washback on teaching materials. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 73-96). Mahwah, NJ: Lawrence Erlbaum.
- Shepard, L. A. (1990). Inflated test score gains: Is the problem old norms or teaching the test? *Educational Measurement: Issues and Practice*, 9(3), 15-22.
- Shih, C. (2007). A new washback model of students' learning. *The Canadian Modern Language Review*, 64(1), 135-161.
- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *Modern Language Journal*, 76(4), 513-521.
- Shohamy, E. (1993). The power of tests: The impact of language tests on teaching and learning. *NFLC Occasional Papers*. Washington, D.C.: National Foreign Language Center.
- Shohamy, E. (2001). *The power of tests: A critical perspective of the uses of language tests*. London, UK: Longman.
- Shohamy, E., Donitza-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298-317.
- Smith, M. L. (1991). Meaning of test preparation. *American Educational Research Journal*, 28(3), 521-542.
- Song, X. M. (2005, February 23). 大学英语四级考试应否与学位挂钩 [Should CET-4 be tied to diploma?]. *Guangming Daily*. Retrieved from [http://www.gmw.cn/01Gmrb/2005-02/23/content\\_184192.htm](http://www.gmw.cn/01Gmrb/2005-02/23/content_184192.htm)



- Spolsky, B. (1994). The examination of classroom backwash cycle: Some historical cases. In D. Nunan, V. Berry, & R. Berry (Eds.), *Bring about change in language education* (pp. 55-66). Hong Kong: University of Hong Kong, Department of Curricular Studies.
- 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.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory, procedures and techniques*. Newbury Park, CA: Sage.
- 调查显示英语四六级考试成部分学生的心病 [Survey shows that the CET-4 and CET-6 have become students' big burden]. (2005, February 25). *Beijing Youth*. Retrieved from <http://news.sina.com.cn/e/2005-02-25/12575938628.shtml>
- Taylor, L. (2000). Stakeholders in language testing. *Research Notes*, 2, 2-4.
- Vernon, P. E. (1956). *The measurement of abilities*. London, UK: University of London Press.
- Wall, D. (1996). Introducing new tests into traditional systems: Insights from general education and from innovation theory. *Language Testing*, 13(3), 334 - 354.
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System*, 28(4), 499-509.
- Wall, D., & Alderson, J. C. (1993). Examining washback: The Sri Lankan impact study. *Language Testing*, 10(1), 41-69.
- Wang, Y. (2004). 大学英语教学改革要有的放矢——从“聋哑英语”谈起 [Reform of the CET should be targeted: Starting from 'dumb English']. *Sino-US English Teaching*, 7(1), 15-22.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, 13(3), 318-333.
- Watanabe, Y. (2004). Methodology in washback studies, In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp.19-36). Mahwah, NJ: Lawrence Erlbaum.
- Weisberg, H., Krosnick, J.A., & Bowen, B. (1996). *Introduction to survey research, polling and data analysis*. Thousand Oaks, CA: Sage.



- Widdowson, H. G. (2003). *Defining issues in English language teaching*. Oxford, England: Oxford University Press.
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.
- Williamson, M. (1994). The worship of efficiency: Untangling theoretical and practical considerations in writing assessment. *Assessing Writing*, 1(2), 147-173.
- Xue, B., & Wang, D. B. (2006, August 12). 专家指出大学生就业难九大原因 扩招影响仍存在 [Experts analyze 9 reasons for difficulty in job hunting and the impact of college enrollment expansion]. *Information Times*. Retrieved from <http://news.sina.com.cn/c/2006-08-12/025010701124.shtml>
- Yang, H., & Jin, Y. (2001). Score interpretation of CET. *Foreign Language World*, (1), 62-68.
- Yalden, J. (1981). *Communicative language teaching: Principles and practice*. Toronto, Canada: The Ontario Institute for Studies in Education Press.
- Zhang, S. (2003). 教一考分离: 大学英语四六级考试改革的必由之路 [Separation of teaching and testing: Unavoidable choice of the reform of CET-4 and CET-6]. *Foreign Language Teaching and Research*, 99(5), 385-386.
- Zhang, X. (2004). 大学英语测试质疑评论 [Questioning the CET]. *Foreign Language Teaching and Research*, 100(2), 65-69.
- Zheng, Y., & Cheng, L. (2008). College English Test (CET) in China. *Language Testing*, 25(3), 408- 417.