

The Impact of University Entrance Examination's Change on High School English Language Teaching and Learning: Experts' and Teachers' Perspectives

Parvaneh Shayeste-Far¹

Assistant Professor in Teaching English as a Foreign Language, Farhangian Teacher Education University, Tehran, Iran

Gholam Reza Kiany

Associate Professor in Teaching English as a Foreign Language, Tarbiat Modares University, Tehran, Iran

Monireh Norouzi

PhD Candidate in Teaching English as a Foreign Language, Tarbiat Modares University, Tehran, Iran

Received: 02.11.2016

Accepted: 06.06.2017

Extended Abstract

1. Introduction

Testing-based reform has often been used as a lever for introducing desirable changes. This tenet is found in light of the claim that what is targeted to be measured by testing, in particular by tests of higher stakes, would strongly determine and shape what gets taught and learnt at the lower stakes. Clear examples are various University Entrance Examinations (UEEs) which have long appeared to have immense importance for the people and institutions involved. In fact, high-stakes tests such as the UEEs are widely perceived to have the potential to shape curricular teaching and learning. Such a test consequential effect, i.e., 'test washback', has encouraged decision makers and educators to manage for a test-based change to promote the quality of education in general and the quality of learning and teaching in particular. Reform policies as such have also been increasingly recognized by policymakers, educators, and research community in Iran, during the past decade. Such a recognition and appreciation of reform policies consequently led to reform initiatives that aimed at enhancing the quality of high school learning and teaching through reforming the conventional UEEs system of the country. The outcome was the 'Act of Student Admission at Universities (ASAU)' which mandates gradual replacement of the English requirements of the UEEs with high school National English Achievement Tests (NEATs) students have to take over their four high school years. However, the existing evidence shows that success is not necessarily ensured as hoped. Reasons for such a failure were mostly studied after 'implementation' or at the 'completion phase' of a reformed test but not during its 'implementation phase'. Such an early evaluation has been mostly neglected. Taking into account the

1. Corresponding Author: parishayeste@yahoo.com

perceptions and perspectives of the English Language Teaching (ELT) practitioners of the country, the present study aimed to explore the effectiveness of the UEEs program change in improving the ELT context of the Ministry of Education (ME).

2. Methodology

A mixed-methods approach was employed to collect both qualitative and quantitative data. For such a purpose, semi-structured interviews with 16 ELT professors and head-teachers followed by questionnaire data from English language teachers were used. Participants were instructors of the Ministry of Science, Research and Technology (MSRT) and ME. Eight university professors and 8 ME teachers, selected through purposive sampling, participated in the qualitative phase, and 406 English teachers (both males and females, with teaching experience of 5 to 30 years), selected through cluster and convenient sampling procedures, participated in different phases of the study. The interview guide was prepared on the basis of the information obtained from the Act, higher-order policy documents of the country, available literature, and the researchers' experience and knowledge of the two admission systems. The interview guide questions sought interviewees' views about the consequences of the new tests on the ME's English Foreign Language (EFL) teaching and learning. The interviews were tape recorded, transcribed, and then carefully content-analyzed.

Next, quantitative data were collected through the researcher-made 'Questionnaire of Change Consequences-Teachers' version (QCC-Ts)' that was distributed among 406 English teachers. The interview analyses together with a review of the related literature gave more than one hundred categories and behaviors showing high-stakes test impacts (e.g., teachers' considerations of practicing the skills measured by high-stakes tests) which later led to the development of a temporary driven model of change impact stated through the QCC-Ts, with a total of 82 items forming three main scales: Teachers' Planning and Instruction, Learning Objectives, and Fairness in admission. A series of factor analyses, both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were run to substantiate the whole scale.

3. Discussion

3.1. Interviews: Professors and Head-Teachers

Data analysis of qualitative phase resulted into the identification of more than 200 themes that were finally subsumed under 14 categories. The necessity of "*Adequacy of the new admission criteria*", "*fairness in assessment through the new program*", "*Students' quality and depth of knowledge*" as well as considerations about the "*quality of the output of the ME*", and "*Resurgence of the similar coaching for the UEEs*" were the most commonly mentioned areas perceived to be negatively influenced by the NEATs-based program. A relative dissatisfaction was found in the interviewees' views on the final outcome of the NEATs program in terms of its effects on the quality, quantity, rate, and depth of learning.

In teaching area, the participants reported a probability for integrating a balanced teaching and learning of the main skills of the English language in the ELT contexts of high schools. However, the respondents found the consequences positive on “*students’ learning activities*” and “*teachers’ assessment of students’ learning*”. That is, the new program will lead to more student engagement in learning activities, and increased regular assessment by teachers.

3.2. Questionnaire Surveys: English Teachers

The results of the factor analysis of the QCC-Ts revealed a significantly good-fitting model for the constructed scales (CMIN/DF=2.79, RMSEA=0.04, CFI=0.981, GFI=0.984, and a small $\chi^2=42$).

“*Classroom activities and time-arrangement*”, and “*teaching materials*” teachers would use in their classes were perceived as the two areas being most strongly driven by the new high-stakes tests ($X=4.40$, $SD=.96$; $X=4.35$, $SD=.79$ for classroom activities and materials, respectively). For instance, perceived consequences of the NEATs on ‘types and structures of classroom activities’ ‘students’ test-taking skills and strategies’, ‘extra time spent on grammar instruction’, and on ‘particular classroom activities purposefully adopted (e.g., vocab., translation, and structure drills) to increase students’ scores in the NEATs’ were reported by mean scores clustering around 4.5 in the Likert scale that illustrates ‘agreement’ or near ‘complete agreement’ with the NEATs impacts.

Inconsistency between the desired policies and the observed practices was found for teachers’ ‘*methods of teaching*’ and ‘*content of teaching*’ that came up with the mean values of 4.01 and 3.59, respectively. For instance, two-thirds of the participants said that they change their teaching methods in order to ‘develop students’ NEAT-taking skills’ or ‘increase students’ scores in these tests’. These show how measurement-driven program of the ME obliges EFL teachers to change their adopted methods, employed activities, and applied materials in the direction of meeting testing objectives rather than learning objectives.

As to the NEATs consequences on learning-related factors, EFL teachers were almost neutral about the NEATs consequences on learning-related factors, especially with regard to ‘*affective consequences*’ ($2.5 < X < 3.5$). Regarding ‘*Cognitive abilities*’, teachers appeared neutral, too. More than half of the teachers did not completely agree that such a particular consequence of the NEATs would come about. From their perspectives, the impacts of the program might not be substantial for all cognitive aspects of students’ learning.

4. Conclusion

The obtained evidence showed some degrees of incongruity between policy and practice as the EFL practitioners and teachers reported. Policy inference obtained from the practitioners’ views on tests power does not confirm a full compatibility between the idealized and the real consequences. In concrete, perceived lower rate

of learning objectives, a rather neutrally felt effect on curricular planning and instruction, and neutrally perceived fairness, reflect the mismatch observed. Overall, tests consequences relate to tests effectiveness question which is at the core of the conceptions of evaluation. Expanding the findings of this evaluation framework could aid in developing more socially directed policies regarding students' assessment, and illuminate non-productive structures and practices within the new assessment program.

Key Words: High-stakes test reform; NEATs; Student admission program change.

References (in Persian)

- Gholipour, R., & Aghabozorgi, M. (2005). *Gozaresh-e barressi-e pishnevise layehaye gozineshe daneshjoo dar Iran [A report on the article for student selection program in Iran]*. Tehran: The Parliament's Office of Cultural Studies.
- Kia, A., & Bozorgi, K. (2006). *Ezhare nazare karshenasi darbareye tarhe nahveye pazireshe daneshjoo dar daneshgaha va marakeze amoozeshe ali-e keshvar [Comments on the proposal for student admission to higher education universities]*. Tehran: The Parliament's Office of Cultural Studies.
- Kiamanesh, A. R. (2000). *Arzeshyabi-e amoozeshe [Educational evaluation]*. Tehran: Payam-e Noor Publication.
- Ministry of Education of the Islamic Republic of Iran (2010). *Sanade meli-e barnameye darsi [The document of national curriculum]*. Tehran: Ministry of Education.
- The Parliament of the Islamic Republic of Iran (2007). *Ghanoone pazirehesh daneshjoo dar daneshgah va marakeze amoozeshe ali-e keshvar [The act of students admission to universities]*. Tehran: The Parliament.
- The Supreme Council of Cultural Revolution of Islamic Republic of Iran (2009). *Naghsheye jame elimi-e keshvar [Iran's comprehensive science roadmap]*. Tehran: The Supreme Council of Cultural Revolution.

References (in English)

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297.
- Arbuckle, J. L. (2009). *IBM SPSS Amos 18 user's guide*. Chicago, IL: IBM.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Camilli, G. (2006). Test fairness. In R. Brennan (Ed.), *Educational measurement* (4th ed., pp. 221–256). Westport, CT: American Council on Education and Praeger.
- Carless, D. (1997). Managing systematic change: A critical analysis of Hong Kong's target-oriented curriculum initiative. *International Review of Education*, 43(4), 349–366.

- Chapman, D., & Snyder, C. W. (2000). Can high stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development, 20*, 457-474.
- Cheng, L. (1998). Impact of public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation, 24*(3), 279-301.
- Cheng, L. (2002). *Taiwanese junior high school English teachers' perceptions of the washback effect of the basic competence test in English* (Unpublished doctoral dissertation). The Ohio State University, Ohio, USA.
- Cheng, L. (2004). The washback effect of a public examination changes on teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 146-170). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge: Cambridge University Press.
- 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. 3-17). Mahwah, New Jersey: Lawrence Erlbaum.
- Cheng, L., & Sun, Y. (2015). Interpreting the impact of the Ontario Secondary School Literacy Test on second language students within argument-based validation framework. *Language Assessment Quarterly, 12*(1), 50-66.
- 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.
- Cheng, L., Watanabe, Y., & Curtis, A. (2004). *Washback in language testing: Research contexts and methods*. London: Laurence Erlbaum Associates, Publishers.
- Decker, D. M., & Bolt, S. E. (2008). Challenges and opportunities for promoting student achievement through large-scale assessment results: Research, reflections, and future directions. *Assessment for Effective Intervention, 34*(1), 43-51.
- Fredriksen, J., & Collins, A. (1989). A system approach to educational testing. *Educational Researcher, 18*, 27-32.
- Hargreaves, A. Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal, 39*(1), 69-95.
- Henrichsen, L. E. (1989). *Diffusion of innovations in English language teaching: The ELEC effort in Japan 1956-1968*. New York: Greenwood Press.
- Hughes, A. (1988). Introducing a need based test of English language proficiency into an English medium university in Turkey. In A. Hughes (Ed.), *Testing English for university study* (pp. 134-153). London: Modern English Publications.
- Kane, M. (2002). Validating high-stakes testing program. *Educational Measurement: Issues and Practices, 21*(1), 31-41.

- Kane, M. (2011). Validating score interpretations and uses. *Language Testing*, 29(1), 3-7.
- Kennedy, C. (1988). Evaluation of the management of change in ELT projects. *Applied Linguistics*, 9(4), 329-342.
- Kennedy, C., & Kennedy, J. (1996). Teacher attitudes and change implementation. *System*, 24(3), 351-360.
- Kiany, Gh., Shayestefar, P., Ghafar Samar, R., & Akbari (2013). High-rank stakeholders' perspectives on high-stakes university entrance examinations reform: Priorities and problems. *Higher Education*, 65(3), 325-340.
- Kikuchi, K., & Browne, C. M. (2009). English educational policy for high schools in Japan: Ideals vs. reality. *RELC*, 40, 172-191.
- King, J. A. (2005). Participatory evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 291-294). Thousand Oaks, CA: Sage.
- Kornhaber, M. L. (2004). Appropriate and inappropriate forms of testing, assessment, and accountability. *Educational Policy*, 18(1), 45-70.
- Kunnan, A. J. (2004). Regarding language assessment. *Language Assessment Quarterly*, 1(1), 235-253.
- Markee, N. (1990, March 10th). *The diffusion of communicative innovations and classroom culture: An ethnographic study*. Paper presented at the 24th Annual TESOL Convention, San Francisco.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell Publishing.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: American Council on Education and Macmillan.
- Munoz, A. P., & Alvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, 27(1), 33-49.
- Nagy, P. (2000). The three roles of assessment: Gate keeping, accountability, and instructional diagnosis. *Canadian Journal of Education*, 25(4), 262-279.
- Pan, Y-Ch., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching*, 9(1), 108-122.
- Pizzaro, M. A. (2010). Exploring the washback effects of a high-stakes English test on the teaching of English in Spanish upper secondary schools. *Revista Alicantina de Estudios Ingleses*, 23, 149-170.
- Rogers, E. M. (1983). *The diffusion of innovations* (3rd ed.). London: Macmillan.
- Segers, M., & Dochy, F. (2001). New assessment forms in problem-based learning: The value added of students' perspective. *Studies in Higher Education*, 26(3), 327-343.
- Shayestefar, P., Kiany, G., Ghafar Samar, R., & Akbari, R. (2012). Reform in high-stakes university entrance examinations of Iran: A critical base-line study. *International Journal of Humanities*, 19(2), 97-122.
- Shohamy, E. (1993). *The power of tests: The impact of language tests on teaching and learning*. NFLC Occasional Paper. Washington, D.C.: National Foreign Language Center.

- Shohamy, E., Donista-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing, 13*, 298-317.
- Smith, M. L. (1991). Meaning of test preparation. *American Educational Research Journal, 28*, 521-542.
- 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.
- Stoller, F. (1994). The diffusion of innovations in intensive ESL program. *Applied Linguistics, 15*(30), 300-327.
- Strauss, A. & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Wall, D., & Horak, T. (2011). *The impact of changes in the TOEFL exam on teaching in a sample of countries in Europe: Phase 3, The role of the course book. Phase 4, describing change*. TOEFL iBT Research Service Report No. TOEFL RR-11-41, Princeton, NJ: Educational Testing Service.
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System, 28*, 499-509.
- Wall, D., & Alderson, J. C. (1993). Examining washback: The Sri Lankan impact study. *Language Testing, 10*, 41-69.
- Wall, D., & Horak, T. (2008). *The impact of changes in the TOEFL examination on teaching and learning in central and eastern Europe: Phase 2, Coping with change*. TOEFL iBT Research Service Report No. TOEFL Ibt-05, Princeton, NJ: Educational Testing Service.
- Watanabe, Y. (2001). Does the university entrance examination motivate learners? A case study of learner interviews. In Akita Association of English Studies (Eds.), *Trans-equator exchanges: A collection of academic papers in honor of Professor David Ingram* (pp. 100-110). Akita, Japan: Author.
- Winke, P. (2011). Evaluating the validity of a high-stakes ESL test: Why teachers' perceptions matter. *TESOL Quarterly, 45*(4), 628-660.