Washback Effect in Teaching English as an International Language

Jim McKinley and Gene Thompson

Washback concerns the impact of tests and assessments on materials, teachers, and learners. Thirty years of research has shown that testing and assessment have influenced language teaching significantly. Teaching English as an international language (TEIL) is a quickly developing and practical new reality that is beginning to influence language assessment and to challenge the use of L1 speaker norms for English language proficiency tests such as IELTS and TOEFL. While the recent, cumulative knowledge of EIL is making those norms increasingly irrelevant, they still have significant impact on World English (WE) users. In this way EIL, which is not bound to a particular variety of English, allows for communication to occur according to context and is at the forefront of change in assessment. Thus, the washback effect in TEIL is driving assessment, while being influenced by the continued power of international language tests.

Intellectual and Social Context

"Washback effect" refers to the influence that language testing has on curriculum design, teaching practices, and learning behaviors. This influence operates in ways that impact the choices of learners and teachers, for example teachers may teach to a test, or learners might focus on aspects of language learning that are likely to be assessed in their future studies. Typically, the washback effect in language testing is seen as either negative or positive, the latter sometimes referred to as *washforward*. It can be seen as harmful to more fluid approaches to language education when definitions of language ability are too limited; it can be beneficial when good teaching practices result. Washback can also be positive or negative to the extent that it either supports or obstructs the accomplishment of educational goals. If a test has positive washback, teaching the curriculum becomes the same as teaching to the test. Negative washback occurs when there may be a mismatch between the stated goals of instruction and the focus of assessment; it may lead to the abandonment of instructional goals in favor of test preparation.

Washback in Language Assessment

The effect of a test on learning and teaching is a concept discussed as early as the 19th century. Research into washback can be traced back to the early 1980s, when the

influence of tests on teaching and learning was first seen as a potential source of bias due to the accountability of test feedback loops. As the results of tests became more important to students (gatekeepers to future prospects), teachers (evaluation), schools (funding), and states (lawsuits), test preparation as a function of teaching became essential. Tests were made to be economical, using multiple-choice questions and focusing on psychometric validity, but perhaps not measuring more complex abilities. Schools and teachers were accountable for student test performance, and thus focused on the skills and outcomes that the tests measured. Given the dynamic interaction between testing and education, the term *systematic validity* was used to refer to the ways in which a test leads to changes in instruction intended to develop cognitive skills that are being measured by a test.

During the 1990s, the research focused on the direction of washback, specifically exploring how tests influence teaching and learning, how tests change attitudes, how tests are used, and the extent to which washback influences different stakeholders. A number of researchers considered the effect of tests on learning and teaching separately from the wider *impact* of tests on classrooms, society, and educational systems. However, this entry considers both *washback* and *impact* in relation to the theory of *consequential validity* in language assessment, in which washback is viewed as a result of testing. Accordingly, *impact* and *washback* are used interchangeably in this entry to refer to the influence of assessment on learning, teaching, and the wider community.

Research has shown the variable extent to which washback influences different individuals in different ways, and the difficulty of *targeting* washback. Significant variability has been noted in the ways that teachers respond to test changes and classroom assessments. Effects may be superficial, indirect, and unpredictable due to individual differences in the way that learning is mediated by teachers, textbook writers, and publishers (Andrews, Fullilove, & Wong, 2002). With globalization, the world has witnessed an increase in internationalization of higher education, which has resulted in a dramatic increase in the number of international students in the last 25 years. The prominence of English alongside this internationalization process has also seen the use of international tests of English such as Test of English for International Communication (TOEIC), Test of English as a Foreign Language (TOEFL), and International English Language Testing System (IELTS) as standard tools in the wider learning community. The increasing weight of these tools in education raised questions about the impact of such tests on teaching and learning, with suggestions that language skills were suffering due to the impact of tests.

Recently, different questions about the impact of tests are being asked. Work from the field of World Englishes has called into question the validity of tests in light of the use of English as an "international" language—specifically the extent to which current tests and test design practices represent the way English is *now* used in the global context (Widdowson, 2012). Accordingly, the washback effect of such tests on learners, teaching, and the wider community has been questioned in the EIL context.

Testing in the TEIL Context

Traditional English language teaching approaches to testing rest on a monolithic view of English, based on the norms of "standard English" derived from inner-circle L1 speakers. Constructs to be tested are defined in relation to models of communicative language ability that are understood as concepts of an ideal monolingual L1 speaker's communicative ability. Alternatively, important considerations about testing within the teaching English as an international language (TEIL) context have been raised, in particular that "recent debates about assessment of English as an international language have revolved around two important questions: Whose norms should we apply? How do we define proficiency in the English language?" (Canagarajah, 2006, p. 229).

In any language testing, criteria designations are crucial. These tests are designed to reveal learners' proficiency levels and progress made in language courses. The standards set in a language test function as constructs that determine the scope and nature of the test in consideration of what levels learners are expected to achieve. It is the definitions of both these standards, as well as achievement, that therefore need to be carefully established. Initially achievement was measured according to L1 speakers of English, specifically either standard American English or British English. However, as more recognition has been given to the reality that there are more L2 users of English, the focus has shifted to TEIL. This has challenged the traditional constructs of English language testing regarding proficiency as they are expected to reflect more the nature of communication as understood in EIL than the ability to match accuracy levels of L1 users.

In standardized language testing, the Common European Framework of Reference (CEFR) provides the best example of a clear use of standards. The CEFR is used throughout the world as the clarity of the descriptions of levels in the framework help to alleviate some of the problems associated with testing constructs. However, the essential construct of EIL is not a consideration in the CEFR. The "can do" statements indicate that proficient users are responsible for producing L1-level language, without consideration of the competence of interlocutors. This is because the expectation is that the interlocutor is an L1 user of English. However, as the majority of English users are EIL users, while those L1 English norms are the boundaries in which shades of meaning are to be contained, there needs to be a cooperative negotiation between communicators in establishing understanding.

The main problem with maintaining test constructs within L1 English norms is that they do not recognize the requirements of EIL, such as accommodating different varieties of English. Furthermore, the reality is that L1 English users are not necessarily capable of proficient EIL communication. The constructs instead need to be redesigned to take into consideration the necessities of an English language user to be able to resolve communication problems. However, articulating EIL constructs so that they can serve as standards for assessment presents significant challenges.

As EIL research has been more sociolinguistic, establishing a learning dimension of EIL communication that clarifies the degrees of competence required of assessment is necessary. Furthermore, an understanding of how learners can progress or improve their language proficiency within this learning dimension needs to be clarified. The significance of these challenges is manifested in the realization that the flexibility required in EIL is dependent on language users' ability to successfully take on a myriad

of different roles dependent on the situation—from everyday workplace communications to subtly nuanced advanced-level discussions. By taking into consideration these necessary flexibilities, the linguistic resources such as grammar and vocabulary, as well as professional knowledge and character, language assessments within the EIL context can be designed to guide learning and establish a better focus on language proficiency. Of course, the challenge is how to specify such criteria.

Within the EIL context, current language testing is too narrow to assess a learner's ability to change, grow, and achieve improved proficiency for negotiation of meaning. As EIL continues to become better established in the research, it is now a critical point in which to articulate EIL constructs in testing. Proficiency as an EIL user, rather than a L1 user, requires redefinition of revolutionary proportion. This involves redefining linguistic norms, reconsidering testing cultures, evaluating the localized validity of test and item design, and considering testing processes (see Brown, 2014 for discussion). It is crucial that educators and education policy makers recognize this development and see proficient English language users as flexible, accommodating, and able to strategically solve problems that arise in communication.

Thus, testing in the TEIL context takes a "plurilithic" view of English that rejects the monolithic "L1-speaker" view of language. It embraces the reality that English is not learned monolingually, but rather bi- or multilingually (see Widdowson, 2012). The particular construct that TEIL provokes is the relationship between competence and performance. By not conforming to L1 speaker norms, L2 users of EIL are assessed as incompetent. Regarding learning achievement of language in use, any nonconformity is also seen as incompetent. However, research in TEIL has made it clear that what may be assessed as "incompetence" in no way impedes EIL users from performing competently in communication (see Widdowson, 2012).

Accordingly, the basis of assessment is set to evolve. In maintaining the principles that underlie language testing, such as construct validity, fairness, and accountability, the testing community has taken a conservative stance. Of course, it is paramount to note in consideration of changes to assessment that EIL norms are constantly in flux and therefore do not serve well as a basis for test design. Making changes to a test based on a wide variety of possible "correct answers" dependent on context, would see test takers unsure of both how to prepare for such a test as well as the standards by which their performance would be judged. In response to this conundrum, it has been suggested that the solution may lie in explaining clearly, for every test, both the purpose as well as the target English(es) (Brown, 2014). For example, as the TOEFL test is based on North American academic English, it should be identified as such, rather than as a test attempting to assess overall English proficiency. A further suggestion is that global definitions are needed for international standardized proficiency tests, and that the clearer definitions provide a defense for test validity (Brown, 2014).

Major Dimensions of the Washback Effect in TEIL

Regarding the direct impact of testing, both learner washback and program washback need to be considered. The distinction is important here, as learner washback is concerned with learners' acceptance (or resistance) of a test's fair assessment, while

4

program washback is concerned with teachers and all other parties involved in influencing and developing courses and materials. The major dimensions of washback include the learners, teachers, materials, and wider impact of the test. The possible impact of different participants can fluctuate greatly depending on the context, and there are often more multifaceted connections between them as contributors to the educational process.

Learners

Of all stakeholders, learners are probably the most affected by washback. However, the influence on learners is most likely mediated by the effect of the test on their teachers, textbook developers, and schools. Preparing for and taking the test, as well as resulting outcomes, will affect the test takers. The effect of tests and test changes can be contradictory. An example is a study of the washback on learning by introducing a speaking test in Hong Kong. Due to the decisions made about them based on the required skills needed in the test, there was evidence of positive washback on learners as they focused on speaking and oral communication (Andrews, Fullilove, & Wong, 2002). However, in preparation for the test, the study also showed that student preparation may have relied on superficial memorization of phrases for the test (i.e., negative washback).

While language learners may be the stakeholders most influenced by language testing washback, studies have focused more on investigating learner perceptions and attitudes, with a few observational studies detailing learner perspective or learner behavior linked with pre- and posttest washback (see, e.g., Green, 2006). Results are conflicting about the effect of test preparation on student test scores. The effects on learners appear to be more individually mediated via motivation, perception and understanding of the test, and prior experiences.

Teachers

Teachers are also affected by washback in terms of the impact that testing has on curriculum design, pedagogical practices, and other facets of language teaching, and are most often the focus of studies of the washback effect. Such studies of the effect of tests on teachers show the results are complex and mediated by different factors. A study of the impact of university exams in Japan explained that both positive and negative washback was noted in observations of five teachers' regular and exam-oriented classes (Watanabe, 2004). While the entrance tests made clear *what* to teach, it did not necessarily affect *how* teachers taught. Therefore, the direct impact of tests on methodology can be contradictory—some teachers focused on grammar translation due to test preparation (i.e., a factor influencing negative washback) while others used test preparation materials to encourage authentic, communicative English. Based on these major differences in the ways that teachers approached the teaching of skills, the author suggested that teachers' *perceptions*—of test content, of their students' skills, and of learners' proficiency levels—were key. Other studies have noted the influence of teachers' beliefs and professional experiences as having more impact on teaching than washback effects from tests (Green, 2006).

Wider factors, such as school culture, also influence the way that exams influenced teaching. In the Japanese context, extant research has shown that teacher beliefs about entrance tests continue to encourage a grammar-translation methodology focused on exam preparation, reinforced by a culture of schools where teacher collaboration worked against innovation. Such teacher beliefs and school systems then affect student knowledge and perceptions of tests. In a study with international students in New Zealand, the relationship between the influence of tests and teachers' knowledge about the assessments showed that different teachers had different amounts of knowledge about the IELTS test, suggesting this as a factor influencing the washback of the test on teaching (Hayes & Read, 2004). Furthermore, the impact of *changes* to tests can take considerable time to affect teacher practices, simply because teachers may not be aware of the ways in which the test has changed. Accordingly, the ways in which tests influence teaching is complex —different teachers respond to tests in different ways.

Materials

From a critical perspective, international language tests act as tools for preserving language standardization and consistency absent of consideration of the realities of how languages are being used in modern, multilingual societies. A key part of that process is the influence of tests on materials. Tests can influence materials negatively when the textbook is the result of washback effect targeting unrealistic language use (Saville & Hawkey, 2004). The impact (and power) of textbook writers and publishers was noted in a study with teachers in Hong Kong who responded to a new speaking test by using published materials that were focused on the specific skills and format needed for the exam (Andrews, Fullilove, & Wong, 2002). Teachers supplemented these materials by encouraging students to memorize word lists of common phrases. Such narrowing of instruction in terms of teaching and materials has been noted in other studies (Green, 2006) and represents ways in which materials conform students to norms—even more than in general instruction. In the World Englishes context (see Brown, 2014), these norms have been challenged as needing to adjust to allow for materials and activities development based on local and international situations. The situations should reflect students' everyday lives, and embrace both L1 speaker and L2 speaker interactions.

The current understanding of testing in a TEIL context suggests the washback effect has shown that maintaining inner-circle, L1-speaker norms is a problem. It is a hindrance to incorporating local varieties in the development of teaching materials and approaches, blocking the growth of such varieties (Canagarajah, 2006).

Test impact and TEIL

In more international approaches to language education, such as that found in TEIL, there is much cause for exploring potential positive and negative washback in teaching and assessment. In the expanding circle, it is necessary to consider the importance of international language tests as gatekeepers to further study and employment. Taking such tests imposes significant monetary costs on learners, but they are seen as a

gateway to studying overseas and therefore strategic in reaching personal and professional goals and opportunities (Moore, Stroupe, & Mahony, 2012).

Also important to consider is that tests impact upon the choices made by parents for their children. An outline was made of the perceptions of parents, learners, and teachers in Hong Kong about international tests of English for young learners marketed globally (Chik & Besser, 2011). Such international tests are being reappropriated by some stakeholders in the Hong Kong context to act as advancement tests for local students to enter better high schools. Furthermore, the financial constraints of taking and preparing for such tests imposes social costs on local parents, and the media adds pressure by portraying such tests as being steps toward further academic success. From the social and financial impact of tests and test preparation, it has been argued that there is unequal access to language testing, which sustains inequality in education systems (Chik & Besser, 2011), suggesting that there are specific effects of tests on speakers of EIL.

Changes in the Washback Effect in TEIL over Time and how it has been Treated

Some have argued that international tests are making more of an effort to respond to global needs by incorporating feedback from stakeholders in English as a foreign language (EFL) contexts. However, initiatives to introduce language tests designed for local learners and users have stalled, as the resistance towards the National English Ability Test (NEAT) in Korea shows. Overall, the debate about norms and criteria continues with respect to large-scale language tests.

One such large-scale language test that has received considerable attention regarding its impact on stakeholders is the UK/Australia IELTS. More than just an exam, this testing system takes a task-based approach using authentic text and discourse that target the assessment of specific language skills, divided into four modules—listening, reading, writing, and speaking. IELTS was first administered in 1980 as ELTS, as it was used only in the UK. It became "international" when the International Development Program of Australian Universities and Colleges was brought on board for the first revision of the test, and the first IELTS test was adminstered in 1989. Since then, some important revisions have been made: in 1995 to better streamline the reading and writing modules; in 2001 for the speaking module; and in 2005 for new writing assessment criteria. What is notable about the IELTS assessment band scale is that it differs from the CEFR's "can do" scales by purposely leaving out any mention of L1 or L1-like use of English. The focus instead is on an "operational" command of English. While this suggests at least some awareness of TEIL, such intentions are not made explicit.

With any major, large-scale testing there will be a number of different stakeholders involved. Stakeholders in IELTS have been identified as insiders, that is, students and instructors, and outsiders, including administrative boards, parents, and employers, who form the language-testing *constituency*. Research on the impact of IELTS has revealed important insights, in particular the IELTS Impact Study (IIS), conducted by Cambridge ESOL examiners, which investigated washback on teaching materials by exploring the roles, responsibilities, and attitudes of the constituency

stakeholders (Saville & Hawkey, 2004). The focus in the IIS on impact in its broadest sense was in response to the limited scope of washback effects on teaching and learning; *impact* in the IIS explored the social consequences of testing. Significant in the study's findings was that ethical language testing is possible with the help of test impact and washback studies. These studies have assisted already by pointing out the critical language testing view that tests are powerful and controlling in considerations of policy changes. Based on this understanding, the IIS takes into consideration a wide range of factors in order to make the distinction between the washback effects on language learning materials and classroom activities, as well as each of the different stakeholders involved.

In a multimethod case study investigating the impact of IELTS on a country, researchers outlined how IELTS affected the entire English-using community. It was noted in the study that a Cambodian variety of English is emerging, and that "it is almost certainly the case that Cambodian test-takers modified their colloquial English to be more aligned with standard dialects for the purposes of the IELTS Test" (Moore, Stroupe, & Mahony, 2012, p. 59). The authors recognized that the primary purpose of the test is to assess academic ability for studying abroad, and found no other use in their study. However, they suggested that local use of the test (such as test scores for employment) would have negative impact for Cambodians proficient in the local variety of English. Finally, a number of stakeholder groups in the study indicated that the IELTS test was too "Eurocentric," where achieving higher scores involved not only language proficiency, but also Eurocentric world knowledge.

Research on the impact of IELTS on students has also highlighted the value of prioritizing student *identities* and *orientation* as they negotiate the challenges they face in their English-medium studies. Significant in the research are students' experiences preparing for and taking the IELTS, as well as students' understanding of the impact of IELTS results both before and during their English-medium studies. This research sees IELTS as a learning and admissions management tool.

In addition to studies on IELTS, washback studies have also focused on the impact of other international tests such as TOEIC and TOEFL. Studies have highlighted the importance of the social consequences of testing, and included, while not always specifically mentioned, is the impact of TEIL.

Current Emphases on the Washback Effect in TEIL in Research and Theory

For more than a decade, increased focus on TEIL has led to some changes in language assessment, at least as it has been researched in inner-circle contexts. For example, a doctoral research study in the UK (Saville, 2009) explored how perceptions are influenced by testing; students were given a detailed questionnaire that included the question, "What are your opinions about learning English?" Responses focused on the students' reasons for learning the language, with the most common reason being related to English as an international language. This is indicative of an important shift toward TEIL.

The impact of changes to classroom-based assessment provides encouraging news for TEIL. For example, there has been a shift of focus onto procedural knowledge

approaches in the investigation of washback (see Canagarajah, 2006). It had previously been argued that because standards can vary and because the varieties of English are considered to be similar, it simply made more sense to base existing tests on inner-circle norms (see Davies, 2009). However, it is important to note that basing tests on inner-circle norms will likely hinder the growth of local varieties as such an approach also hinders the growth of teaching materials and methods for these varieties. And while context-based tests are central, they essentially ignore current understandings of postmodern communication, which is increasingly based in the realm of World Englishes (see Brown, 2014). "Local" is no longer a single variety, and "international" does not refer to a single best variety, which can be seen in the shift toward English as a lingua franca (ELF) (see McNamara, 2014) a shift shown in studies where English is used as such (Kim & Elder, 2015). Thus, it is important to target receptive skills in the inner-circle varieties of local communities, an achievement possible through TEIL.

World Englishes, English as a Lingua Franca, and TEIL

It has become clear that an understanding of World Englishes is important for inner-circle speakers, as "diversity has become a marker of postmodern life, and technology and globalization have increased access to [World Englishes] users and their writing" (Wetzl, 2013, p. 207). In a mixed-method study with an experimental design, the use of a World Englishes approach to composition course was investigated at an American university (Wetzl, 2013). Surveys, interviews, and essays were used to assess experience to World Englishes and changes in student perceptions. The study found that there was a positive impact of exposure to World Englishes texts and learning about linguistic diversity in the class. It was argued that this led American English speakers to shift their perceptions of World Englishes towards linguistic acceptance. The study suggests ways in which washback can be achieved via changes in curricula, and highlights a more important factor: In a TEIL context, what should be the response to the way in which L1 speakers interact and negotiate meaning in a global world? Is it time to judge *their* capabilities to interact effectively in English as a lingua franca communication?

This issue was highlighted in a study that showed that communication difficulties between air-traffic controllers and pilots from inner and outer circles were often due to unnecessarily complex language use by the inner-circle speakers (Kim & Elder, 2015). While it is understood that not all L1 speakers are proficient English as a lingua franca users, policies are set so L1 speakers are not assessed (McNamara, 2014). Such studies as Kim and Elder's (2015) signal the vital need for recognition of English as an international language used in a global context and suggest ways that language testing needs to reflect such a change by being more specifically focused on the target domains of international use.

Future Directions in Research, Theory, and Methodology

There continue to be some significant gaps in EIL research related to assessment. While the 1970s and 1980s involved the development of theories of communicative

competence, it was not until the 1990s that a systematic model of language ability was developed, which influenced test development. Accordingly, it can be expected that there must be a period of talking about change, before change occurs. Time is needed before the implications from EIL (and language education generally) concerning the authenticity and usefulness of tests, the social nature of negotiated meaning (and how to measure that psychometrically), and how to define English as a medium of communication between L2 speakers is integrated into testing. Examples from in-class assessment, which develops and responds more quickly than testing, show that the washback effect of TEIL is already influencing the field in unexpected ways, such as the need to test L1 speakers' ELF competence.

The understanding of TEIL in applied linguistics indicates that in the near future language testing and assessment will certainly be affected by the rise of English as a global language. TEIL is a quickly developing and practical new reality. While L1-speaker norms have maintained their influence in terms of the infamous high-stakes English language proficiency TOEFL and IELTS, the recent, cumulative knowledge of EIL is making those L1-speaker norms increasingly irrelevant. With the increase of international trade and the development of political multinationals such the European Union, English is often used for communication that involves no L1 speakers of English. In this way EIL, which is not bound to a particular variety of English, allows for communication to occur according to context and free from any restraints presented by geography or ethnicity. Because EIL involves varieties influenced by L2 English contexts, it is more subject to reform.

As EIL continues to become more stable through the establishment of clearer features of context-driven English, it is inevitable that there will be significant debate on language norms and proficiency regarding teaching and assessment. The field of applied linguistics should expect to see more specific discussion on the recognition of the way language norms are influenced by use and context. These "new norms" will need to be designed to accommodate EIL as well as local Englishes in more contextualized versions of English language testing. As the world continues to diversify its peoples' ethnolinguistic backgrounds, language testing will need to account for EIL as well as local varieties of English. With this understanding, students need different strategies so they can transfer between different populations, as well as the local and the global. Both norms and codes will vary in each of these different contexts. Davies (2009) noted that the global use of Standard English serves many purposes, and the necessity of standards for formal uses of English for assessment purposes is what maintains the norms of standard Englishes (i.e., British or American) over World Englishes. But the argument for upholding a universal English proficiency for assessment purposes continues to wane.

The changes in the way that English is used as a global language suggest that there are many areas where the intersection between testing and TEIL can be developed to influence positive washback. Few studies have examined the impact of international tests from an EIL perspective. The consequences of tests are *real*. Testing needs to reflect the contexts of professionals working in international contexts, such as air-traffic controllers, and of other stakeholders affected by the norms, proficiencies, and procedures used. Although the impact of TEIL *in* testing has yet to be seen, the impact of TEIL *on* testing is coming.

SEE ALSO: Assessment Practices in Teaching English as an International Language; Challenges Faced by Non-Native English Speaker Teachers; Discrimination and Discriminatory Practices Against Non-Native English Speaker Teachers; English for Specific Purposes; Intensive English Language Programs; Textbooks; Ownership of English and Language

References

Andrews, S., Fullilove, J., & Wong,Y. (2002). Targeting washback—a case study. *System, 30*(2), 207–23. doi: 10.1016/S0346-251X(02)00005-2

Brown, J. D. (2014). The future of World Englishes in language testing. Language Assessment Quarterly, 11(1), 5–26. doi: 10.1080/15434303.2013.869817

Canagarajah, A. S. (2006). Changing communicative needs, revised assessment objectives: Testing English as an International Language. *Language Assessment Quarterly, 3*(3), 229–42. doi: 10.1207/s15434311laq0303_1

Chik, A., & Besser, S. (2011). International language test taking among young learners: A Hong Kong case study. Language Assessment Quarterly, 8(1), 73–91. doi: 10.1080/15434303.2010.537417

Davies, A. (2009). Assessing World Englishes. Annual Review of Applied Linguistics, 29, 80–9. doi: 10.1017/S0267190509090072

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(4), 333–68. doi: 10.1080/15434300701333152

Hayes, B., & Read, J. (2004). IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In L. W. Cheng, Y. Watanabe, & Curtis, A. (Eds.), Washback in language testing: Research contexts and methods (pp. 97–111). Mahwah, NJ: Lawrence Erlbaum Associates.

Kim, H., & Elder, C. (2015). Interrogating the construct of aviation English: Feedback from test takers in Korea. Language Testing, 32(2), 129–49. doi: 10.1177/0265532214544394

McNamara, T. (2014). 30 Years on—evolution or revolution? Language Assessment Quarterly, 11(2), 226–32. doi: 10.1080/15434303.2014.895830

Moore, S., Stroupe, R., & Mahony, P. (2012). Perceptions of IELTS in Cambodia: A case study of test impact in a small developing country. IELTS research reports, 13, 1–109.

Saville, N. D. (2009). Developing a model for investigating the impact of language assessment within educational contexts by a public examination provider. Unpublished doctoral thesis, University of Bedfordshire.

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.

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.

Wetzl, A. M. (2013). World Englishes in the mainstream composition course: Undergraduate students respond to WE writing. Research in the Teaching of English, 48(2), 204–27.

Widdowson, H. G. (2012). ELF and the inconvenience of established concepts. *Journal of English as a Lingua Franca, 1*(1), 5–26. doi: 10.1515/jelf-2012-0002