Preparing for Admissions Tests in English

1. Introduction

Test preparation for admissions to education programmes has always been a contentious issue (Anastasi, 1981; Crocker, 2003; Messick, 1982; Powers, 2012). For Crocker (2006), “No activity in educational assessment raises more instructional, ethical, and validity issues than preparation for large-scale, high-stakes tests.” (p.115). Debate has often centred around the effectiveness of preparation and how it affects the validity of test score interpretations; equity and fairness of access to opportunity; and impacts on learning and teaching (Yu et al., 2017). A focus has often been preparation for tests originally designed for domestic students, for example SATs (e.g., Alderman & Powers, 1980; Appelrouth et al., 2017; Montgomery & Lilly, 2012; Powers, 1993; Powers & Rock, 1999; Sesnowitz et al., 1982) and state-wide tests (e.g., Firestone et al., 2004; Jäger et al., 2012), but the increasing internationalisation of higher education has added a new dimension. To enrol in higher education programmes which use English as the medium of instruction, increasing numbers of international students whose first language is not English are now taking English language tests, or academic specialist tests administered in English, or both. The papers in this special issue concern how students prepare for these tests and the roles in this process of the tests themselves and of the organisations that provide them.

2. International higher education and the role of admissions tests in English

In the context of expanding participation in higher education and budgetary constraints, universities in majority English speaking countries such as the USA, UK, Australia, and New Zealand have increasingly come to rely on internationalisation as the key to income growth. This has continued to involve attracting international students to domestic campuses. The number of internationally mobile tertiary students has grown by almost 5% per year over the past two decades (reaching 5.6 million in 2018) and the number of international students in these four countries is now over ten times greater than the number of students from them studying elsewhere (Organization for Economic Cooperation and Development (OECD), 2020). More recently, there has been a trend for higher education institutions in such countries to open overseas campuses or establish partnerships in territories in or close to the major sources of their international students (notably China and other parts of East Asia). The expansion of such transnational education means, for example, that there are now more international students studying for UK accredited degrees outside the UK than within UK borders (Higher Education Statistics Agency (HESA), 2020). Conceptions of education as a public benefit reflected in policies that encourage widening domestic access to higher education sit uneasily alongside conceptions of education as a lucrative global commodity reflected in economically and linguistically selective policies for international students.
The privileged status of English as the global language of commerce makes it an important source of cultural capital and its use as the medium of instruction is one of the attractions of an internationalised education. As a result, the growing number of higher education internationalisation initiatives in non-English-speaking countries, designed to attract more students from other countries, have often entailed the introduction of English-medium courses. Reflecting the function of tests as a device for selection, standards for entry often reflect as much the relative standing of the receiving institution in global rankings as any proficiency threshold needed to enable academic success. Tests of English such as the Anglo-Australian IELTS have accordingly witnessed dramatic growth. When it took its current form in 1995, the annual candidature for IELTS was around 45,000, but by 2019 this had surged to 3.75 million (www.ielts.org). Because of the prominence of language skills, it is not surprising that most of the large-scale international tests addressed by the papers in this special issue are concerned with language abilities.

One effect of the international marketisation of higher education has been to encourage dramatic growth in the use of international tests as gatekeeping devices. Little is done to accommodate to the experiences of the increasingly international student population. Established curricula, developed for domestic students, are not often or much adjusted to meet the needs of these students in relation to pedagogy, content or language. International students are instead expected to win access by demonstrating their readiness for academic study on the institutions’ terms through test scores. Although these are most often tests of English language abilities, a small, but growing number of international students are also required to take specialist admissions tests originally designed for first language English users. Examples include the BioMedical Admissions Test (BMAT), the Mathematics Admissions Test, and the Thinking Skills Assessment. These make varying demands on test takers’ English language proficiency. Just as domestic national matriculation tests may dominate the later stages of secondary education, international tests have come to play a central role in how prospective international students prepare themselves for their academic studies.

3. Test preparation: an overview

The use of admissions testing to regulate entry has encouraged the growth of a global ‘shadow education’ industry, operating alongside formal schooling (Bray, 2007), that prepares prospective students to take these tests. For example, as Ross (2008) noted, test preparation in some Asian nations (notably China, Japan and South Korea) has become a “massive enterprise which can exert considerable influence against assessment modernization when it comes into conflict with the vested interests of the cram school industry” (p. 7). Like tertiary study itself, test preparation courses are available both in students’ home countries and in English speaking destination countries. The cost of test preparation can be substantial, adding to the already considerable expense of accessing international higher education and increasing the financial pressures on prospective students and their support networks. These further limit opportunities, especially for those with the least access to effective language instruction at school.

Where there is an admissions test, test preparation will also be found; and it comes in many guises (Latham, 1877). Various terms have been used to describe these: for example, coaching,
teaching to the test, individual tutoring, test familiarization, measurement-driven instruction, and test-wiseness or test-taking strategy training. In this editorial and special issue, we use test preparation as an all-encompassing and neutral term to refer to what Messick describes as “any intervention procedure specifically undertaken to improve test scores, whether by improving the skills measured by the test or by improving the skills for taking the test, or both” (Messick, 1982, p.70). The industry is not limited to cram schools, but also comprises private tutoring, publications and online resources to support self-directed study with producers ranging from individuals to multinational enterprises. Learning processes are diverse: test preparation “may fall anywhere in the broad range bounded by the two extremes of practice and instruction, embracing any combination of test familiarization, drill-and-practice with feedback, motivational enhancement, training in strategies for specific item formats and for general test taking (including advice on pacing, guessing, and managing test anxiety), subject-matter tuition and review, and skill-development exercises.” (Messick, 1982, p. 70).

With appropriateness and ethicality of test preparation as the key factors to consider, most debate on test preparation focuses on three fundamental measurement issues (Crocker, 2006): the validity and defensibility of inferences and interpretations drawn from test scores when test performance is affected by different types and amounts of test preparation; the consequences of test use when time and resources are dedicated to test preparation; and the fairness of testing regimes when some test-takers have “differential access” to costly test preparation programmes and materials.

Research studies on preparation for admissions tests for international higher education such as GRE (Powers, 1985a, 1985b, 1987), TOEFL (e.g., Alderson & Hamp-Lyons, 1996; Hamp-Lyons, 1998; Ling et al., 2014; Liu, 2014; Matoush & Fu, 2012; Yu et al., 2017), IELTS (e.g., Gan, 2009; Green, 2006, 2007; Read & Hayes, 2003; Saville & Hawkey, 2004; Winke & Lim, 2017), and PTE-A (e.g., Knoch et al., 2020; Wei, 2017) can be classified broadly into two main strands: one concerns the processes and immediate effects of test preparation, the other its effectiveness and products.

Research on the process of test preparation investigates, for example, how participants (usually teachers and test-takers) prepare inside or outside the classroom, what materials they use to prepare for tests, how features of tests and test design affect test preparation, what strategies they practise to manage their time and anxiety, how they maintain or boost their learning motivation and set their goals, their attitudes and beliefs about test preparation, how they manage multiple repeated test-taking (see, for example, Green & van Moere, 2020), and to what extent there exist equity and access issues in test preparation. Research on the effectiveness and products of test preparation investigates the outcomes, value or benefit of test preparation: score gains, improvement of language proficiency, success in future academic study (e.g., Clark & Yu, 2020).

Evaluation of products may involve multiple stake-holders (including teachers and test-takers) and the contribution of “mediating and moderating factors” (e.g., gender, age, socio-economic status) besides test-takers’ language and academic abilities. As for the effectiveness of test preparation, “Few clear main effects are found and interactions abound” because the main effects of test preparation “depend on the outcome measured, the preparation method used and the population tested” (Bishop & Davis-Becker, 2016, p.558), among many other mediating
and moderating factors. One example of this strand of research is Yu et al. (2017), which investigated both the processes and effectiveness of intensive courses in China preparing students for the TOEFL iBT speaking test. The study confirmed that test preparation was a “hugely complex, multiple-component construct” (p.1). The teaching of test taking strategies was the most prominent feature of the intensive preparation courses. Teachers and test takers agreed on the usefulness of various test preparation activities. Differences among test takers (n>1500) in their focuses and test preparation activities were significantly related to their age, although effect sizes were small. The researchers also reported only a weak correlation between test preparation activities and test performance.

Opinions on test preparation vary widely: some researchers believe that test preparation improves the validity of test score interpretations by, for example, familiarizing test takers with test formats. Others view test preparation negatively: a potential threat to the validity of test score interpretations and a source of construct-irrelevant variance (Bishop & Davis-Becker, 2016). These debates often come down to judgements about which test preparation practices are considered acceptable and which are considered inappropriate, unacceptable, unethical, or even fraudulent (Mehrens & Kaminski, 1989; Moore, 1994; Popham, 1991).

4. The papers: a short summary

Three of the four papers in this special issue come from test providers and present their perspectives on the roles they play in the test preparation marketplace. Reflecting a practical interest in taking responsibility for the consequences of test use, all three describe interventions by the test providers that aim to make test preparation resources more useful and accessible to learners. All three touch on the use of computer technology in delivering materials and enabling wider access. Their initiatives express a shared interest in engaging with, and taking account of the views of teachers and learners and their expectations of test preparation. While these papers focus on attitudes towards products intended to support test preparation. The other paper (Saifa, Ma, May and Cheng) draws attention to limitations on the influence that intervention by a test provider may have on test preparation practices. The test preparation processes they reveal express the ingrained beliefs of learners, teachers and administrators that will inevitably shape the ways in which test provider initiatives can affect their practices.

In the first paper, O’Sullivan, Dunn and Berry note that most of the literature on test preparation for international language tests is contextually limited, usually being conducted within a single country and often at the level of an individual institution with small numbers of participants. There is a surprising lack of larger-scale cross-regional studies that take account of the effects of learner background on test preparation. The authors set out address this gap through a global survey that investigates variations in test preparation preferences among potential students in different educational, political and social contexts across Asia, Africa and Latin America. Their survey respondents ranked twelve common test preparation options identified by the research team (such as accessing timed practice tests or online tutors). The survey questions also included learner background and expectations for the length (time to test date) and intensity (study hours per week) of preparation. A principal components analysis revealed relatively little variation in preferences and the researchers concluded that similar preparation delivery formats and materials are likely to be welcomed by learners across regions. The scope of the study was
restricted, offering no indication of how successful these options might be in improving scores, how they might affect score interpretations, or how they differentiate test preparation from other learning purposes. Nonetheless, the findings suggest that test takers around the world have similar understandings and expectations of the length, intensity and format of test preparation courses. This challenges assumptions that test preparation practices differ substantially across countries with distinct assessment cultures. Clearly much more work is needed to establish how practices vary, but this study suggests that the way forward lies with comparative studies of the phenomenon.

Sharing a cross regional perspective with O’Sullivan, Dunn and Berry, the second paper by Saifa et al. takes a more qualitative approach. Sites in Australia, Iran and China serve as case studies with data derived from multiple sources: classroom observations, interviews and focus-group discussions. The concentration on three institutions facilitates the inclusion of a wider range of participants (administrators in addition to teachers and students) as well as the investigation of classroom processes. Like O’Sullivan et al., the researchers found much in common across the three contexts. Students were highly motivated, teachers were recruited and valued for their expert knowledge of the tests concerned and the courses were all closely aligned to the structure of the target test. Like earlier studies of preparation courses for similar tests carried out in a diversity of settings, such as, for IELTS, Hayes and Read (2004), Saville and Hawkey (2004) and Green (2007), this suggests a common reflex towards test preparation. The default approach takes the test structure and tasks as the basis for any test preparation programme with completion of test or test-like tasks taking up much of the time spent in the classroom. As in many of the earlier studies, teachers reported feeling constrained to focus on the test tasks and to avoid straying beyond them to, for example, address aspects of academic language use not directly involved in the test. Conversely, as in Alderson and Hamp-Lyons (1996), at least some of the students reported feeling that the focus on the test was too narrow.

Saifa et al. identify some notable differences between the EFL settings (China, Iran) and the ESL setting (Australia), but these seem to reflect more the status of English in these countries than different attitudes towards assessment. In the EFL test preparation courses, the teachers were L2 English users who could share experiences both of successful test taking and of international study. In this they could serve as role models for their students, providing a link between the linguistic and cultural settings they shared with their students and the English-dominant settings to which those students sought access. In contrast, the L1 English speaking teachers in Australia, where students regularly encountered English outside the classroom, derived their authority from their status as native speakers.

The third paper by Gu, Davis, Tao and Zechner brings attention to the part that technology can play in increasing access to test preparation. Obtaining guidance on spoken language use can be particularly challenging for students in EFL settings where there are few opportunities to interact in English (see also Yu et al., 2017). Online practice speaking tests offer them a chance to speak in English, to become more familiar with the test taking experience (which involves recording a response to computer-delivered recorded prompts) and (exploiting speech recognition, analysis and automated scoring technologies), to receive feedback in the form of test scores and diagnostic comments. The study explores teacher and student perceptions of the feedback that the automated system can offer and suggests improvements that can make this more useful.
Although both groups welcomed the feedback as helpful both for test preparation and for more general language development, the researchers also acknowledge the limitations of current technology. First, the state of the art in speech evaluation restricts the feedback to matters of delivery (pronunciation and fluency) and language use (grammar, vocabulary), excluding other aspects of the speaking test construct that feature in the rating scales: topic development (organisation, coherence) and content. Second, making effective use of such feedback as the system can provide may require knowledge that learners generally lack. In other words, to be effective as a guide for learning, feedback requires mediation by a trained teacher. The potential for the system to equalise access to test preparation thus remains limited.

Unlike the language tests addressed by the other three papers, which are taken mainly by prospective international students, the BioMedical Admissions Test addressed by McElwee, Cheung, Cromie, Shannon and Gallacher is also taken by domestic students. Faced with growing numbers of international test takers and pressure to increase participation in medical education among under-represented groups, this paper connects debates over the role of testing and of test preparation in internationalisation to its role in widening, or restricting access at the national level. A validation framework originally developed with language tests in mind (Weir, 2005; 2013 socio-cognitive framework) is suggested as a useful corrective to the emphasis on predictive validity in previous research into admissions tests. As well as incorporating the impact of the test and test scores on stakeholders, the framework offers a systematic basis for addressing the needs of test takers in relation to their physical, psychological and experiential characteristics.

As one aspect of this framework, the paper focusses on the efforts made by the test provider to bring about positive consequences through the concept of “impact by design” (Saville, 2012). In relation to test content, the needs of receiving medical schools are balanced by efforts to take account of the diversity of prospective test takers in relation to their linguistic knowledge and their conceptualisation of scientific topics. Beyond the test itself, the provision of a free revision guide intended to familiarise test takers with the test is considered an important contribution to the consequential aspect of its validity, meeting the requirement in the AERA, APA, NCME (2014) Standards for Educational and Psychological Testing that “Information about test content and purposes that is available to any test-taker prior to testing should be available to all test-takers” (p. 133).

5. Discussion and Conclusions

Any attempt to discourage or prohibit test preparation is surely futile: the incentives for teachers and learners to succeed on such tests are simply too great to resist. Instead, we argue in favour of a holistic approach to understanding the dynamic and complex triad of instruction, ethics and validity issues involved in test preparation (Crocker 2006). Improving the current situation must involve multiple stakeholder perspectives: especially those of test takers, teachers, test providers, and score users (in this case those responsible for university admissions policies). All have an essential role to play in achieving more appropriate, ethical and learning-oriented forms of test preparation: forms of preparation that do not only offer practice in completing test tasks, but that engage learners in developing their understanding of how to operate successfully in the settings that scores are intended to generalise to. The papers in this issue exemplify a range of
strategies for taking account of these perspectives from online surveys to focus-group interviews and classroom observation.

The cross regional comparisons made by the first two papers (O’Sullivan et al. and Saifa et al.) raise the possibility that there may be less variation in test preparation across different educational, political and cultural contexts than is sometimes assumed. It is apparent that expectations for the length, intensity and format of preparation courses and of the close relationship between associated materials and test tasks seem to be widely shared among learners, teachers and administrators, although it remains unclear why such a consensus should exist, or why it should take this form. It may be that the commonality across educational, political and cultural contexts reflects general agreement among test providers, test takers and their teachers on the ultimate goal of test preparation: to achieve a grade that guarantees admission. It also seems to reflect an over-optimistic tendency to identify success on international English language tests with possession of the full range of language abilities needed to support academic success, heedless of the diversity of test methods and the concerns about construct representation raised in the literature.

Although similar test preparation practices were observed in different regions, the influences of local educational cultures and practices were also apparent. The use of a common methodology, as in the first two papers of this issue, helps us to better understand such influences. However, it is clear that there remains much to learn about how individuals negotiate test preparation under local circumstances. Building our understanding of why people choose to prepare for tests in the way that they do clearly requires rich descriptions of the social and educational contexts in which they operate. Tests may have common implications for preparation materials and course content, but are only one among the many factors that influence how preparation is realised. It remains true, as Hamp-Lyons (1997) suggested a quarter century ago, that “Many more studies are needed of students’ views and their accounts of the effects on their lives of test preparation, test-taking and the scores they have received on tests” (p. 299).

A focus on individual test takers opens up an array of questions to investigate. These include, for example, how test taker characteristics (e.g., gender, age, first language, language proficiency, motivation, self-regulation, goal-setting) influence decisions to engage in different test preparation activities; why and how some repeatedly re-take (and prepare for) tests over an extended period; the impact of multiple re-takes (and repeated intensive preparation) on performance and their implications for test validation (see Green & van Moere, 2020); and the circumstances under which test preparation leads to memorized or rehearsed responses in speaking and writing tests (Burton, 2020). There is scope to investigate how test takers adapt their behaviour when preparing for different tests directed towards the same admissions purpose (e.g., when preparing for IELTS, TOEFL iBT, or PTE-A); the effects of different task types (e.g., integrated vs. independent speaking tasks (see for example Wei (2017), or of different delivery platforms or modes (e.g., computer-based speaking tasks in TOEFL iBT and PTE-A vs. face-to-face interviews in IELTS).

There now exists a substantial knowledge base concerning the effects of technology use in language test design and delivery, especially in terms of differences and similarities between computer-based and paper-based language tests, and, more recently, between different platforms (e.g., mobile device vs. fixed workstation). Chapelle (2008) argued that “as computer-
assisted language assessment has become a reality, test takers have needed to reorient their test preparation practices to help them prepare for new test items” (p. 127). The use of technology in test preparation, touched on in all the papers in this issue, introduces two important considerations. First, test takers need to be familiar with the relevant technology when taking computer-based tests. Second, there is now a wide range of apps, designed by individuals, test providers, or test preparation companies, to prepare learners for international English language tests, regardless of whether the tests themselves are computer- or paper-based.

Many of these apps claim the use of artificial intelligence to provide instant personalised feedback based on the use of automated evaluation systems. Some have functions to track and manage progress, and some also provide virtual partners for practising speaking tests. This represents a significant change in the availability of tools and materials for test preparation. However, research on the effects of technology use for test preparation is only beginning to emerge. What can technology do to improve the efficiency of test preparation? To what extent can the use of technology change test takers’ preparation preferences and behaviours? To what extent can the use of technology promote learning-oriented forms of test preparation? To what extent can the use of technology level the playing field by increasing equal access to test preparation materials, or rather exacerbate the digital divide in test preparation? As Bishop and Davis-Becker (2016) rightly pointed out, “equity is a consideration in many aspects of test preparation, including access to current technology” (p. 558). While the world bank estimates that less than half of the world’s population has access to the internet (data.worldbank.org), in these papers there is more focus on whether the existing test taking population finds preparation tools useful – less on how the tools increase access or on how this might be gauged.

Technology use can be a double-edged sword. Technology is being used to enhance test security and to detect cheating. On the other hand, it is also used to undermine test security with negative implications for validity. Technology-based test preparation, if used fraudulently, can facilitate cheating and other forms of malpractice, with the potential for these to occur on a much larger scale than they would in any traditional form of face-to-face classroom-based test preparation programme. Test preparation practices, assisted by technology use, can alter or significantly influence the coachability of the test and the extent to which the intended construct of assessment can be realised. It is therefore increasingly important that we include a consideration of test preparation in the construct analysis of any test: the value of a test cannot adequately be judged in isolation from the practices involved in preparing for it.

To minimise large-scale malpractice in test preparation and to address equity and access issues, testing organisations should provide test takers with up-to-date, learning-oriented test preparation materials in a form that is easy to access (see for example Green, 2017). They should offer guidance on what are and what are not appropriate test preparation practices; encourage a range of suitable approaches to test preparation; and provide test preparation teachers and other key stakeholders (including university admission policy makers) with training and professional development opportunities on an ongoing basis (see also Bishop & Davis-Becker 2016). As Xi et al. (2013) suggested, “Because not only populations taking an assessment but also test preparation strategies can change over time, the psychometric characteristics of test items must be continuously monitored, as initial positive results do not guarantee the maintenance of quality over time.” (p. 9).
The strategies that test takers adopt to prepare for admissions tests can change over time. Sometimes, learners themselves may decide to adopt new approaches, for example because they are attracted by new technology, or persuaded that newly encountered preparation methods will be more effective. Repeated failure on one may lead test takers to try their luck with another of the tests accepted by their chosen institution, or persuade them to adopt new preparation tactics. Decisions can also be driven by external events. During the Covid-19 pandemic some test takers switched from preparing for one test to preparing for another simply for reasons of availability. As a result of the closure of in-person test centres, testing organisations that were able to offer tests that could be taken at home over an internet connection gained market share. Sometimes test takers switch in response to changes in policies on language proficiency requirements on the part of governments or universities.

It appears to us that the part played in test preparation choices by university admissions policies, influenced by market considerations, and the assumptions that underlie them merit fuller investigation. It may be that these policies play a more significant role than the test takers’ own characteristics (e.g., language ability, socio-economic status) or the characteristics of the tests concerned (in terms of test format and content). It may also be that the tendency on the part of receiving institutions to rely on a single measure of language abilities inevitably encourages teachers and learners to focus narrowly on the measurement instrument rather than seeking to understand the demands that English-medium study will place on them.

In the competitive and linguistically and economically selective international higher education arena, understanding preparation for admissions tests requires multiple perspectives. Practices should be viewed through the lenses of a wide diversity of stakeholders (especially test takers, teachers, test providers and university admissions policy makers). This will allow us to (a) better understand the dynamics of the processes, effects, ethicality, fairness, equity, social justice, and consequences of test preparation for admissions purposes, and (b) more importantly, to achieve appropriate, ethical, and learning-oriented forms of test preparation.

Disclosure statement
No potential conflict of interest was reported by the authors.

ORCID
Guoxing Yu, https://orcid.org/0000-0001-5997-2734
Anthony Green, https://orcid.org/0000-0003-4893-1798

References


