Revising the BioMedical Admissions Test (BMAT) to improve impact and washback for candidates and support fair access to test preparation

Sarah McElwee, Kevin Y.F. Cheung, Stephen R.T. Cromie, Mark Shannon & Tom Gallacher

To cite this article: Sarah McElwee, Kevin Y.F. Cheung, Stephen R.T. Cromie, Mark Shannon & Tom Gallacher (2019): Revising the BioMedical Admissions Test (BMAT) to improve impact and washback for candidates and support fair access to test preparation, Assessment in Education: Principles, Policy & Practice, DOI: 10.1080/0969594X.2019.1639133

To link to this article: https://doi.org/10.1080/0969594X.2019.1639133

Published online: 22 Jul 2019.

Submit your article to this journal

Article views: 53

View related articles

View Crossmark data
Revising the BioMedical Admissions Test (BMAT) to improve impact and washback for candidates and support fair access to test preparation

Sarah McElwee, Kevin Y.F. Cheung, Stephen R.T. Cromie, Mark Shannon and Tom Gallacher

Cambridge Assessment Admissions Testing, University of Cambridge, Cambridge, UK; Cambridge Assessment English, University of Cambridge, Cambridge, UK

ABSTRACT
The BioMedical Admissions Test (BMAT) has been used to select students for healthcare courses for 15 years. Recently, the candidature has included an increasing number of test takers who did not complete their schooling in the UK. In line with responsibilities to promote widening participation, a revision of the Section 2 Scientific Knowledge and Applications specification was conducted, to support the preparation of candidates from under-represented groups and international candidates. The process was guided by validity frameworks established in language testing research, with a particular focus on consequential validity and washback. Candidate surveys are reported to illustrate the focus on test preparation. Issues identified in the review demonstrate the impact of international test-takers’ linguistic backgrounds and differences in teaching approaches, and possible interactions between these are discussed. It is recommended that admissions test developers make use of holistic frameworks of validity from other contexts to systematically interrogate a wide range of validity evidence relating to their tests.

ARTICLE HISTORY
Received 1 March 2017
Accepted 26 June 2019

KEYWORDS
Medical education; BMAT; validity; test preparation

Introduction

Admission to study medicine and other healthcare disciplines is highly competitive, attracting the most academically able students who submit to a multi-stage selection process. Issues of widening participation and social justice are often central to discussions around selection.

Universities aim to admit candidates with potential to succeed in the profession from a wide range of backgrounds, and the British Medical Association (2009, p. 8) has commented that ‘doctors should be as representative as possible of the society they serve in order to provide the best possible care’. In the UK, the Selecting for Excellence report (Medical Schools Council, 2014, p. 35) referred to a ‘widening participation problem’, because students from fee-paying schools are over-represented amongst...
applicants to medical schools. Admissions tests, as part of the selection process, face particular scrutiny in terms of their fairness and equity for candidates, and the extent to which they provide a level playing field to assess the suitability of students from a range of school backgrounds, socio-economic statuses and nationalities. To support this aim, the availability of high-quality test preparation materials should be equal across groups, to lessen the risk that admissions tests present a barrier to equal access. Particular care is needed to define any curriculum that a test is based upon, to ensure that it is accessible to all prospective test-takers when they are preparing.

Additionally, materials must be accessible for those preparing worldwide. Internationalisation of higher education and increasing student mobility entails greater numbers of overseas students coming to the UK for biomedical study. The number of students taking biomedical courses taught through the medium of English in other countries is also growing. Therefore, it is important to consider the impact of different educational systems on test preparation and performance, as these issues potentially affect the validity of admissions tests.

This paper describes a revision of the Scientific Knowledge and Applications section of the BioMedical Admissions Test (BMAT), which was prompted by a need to ensure that the increasingly diverse candidature had appropriate information to adequately prepare for the test, and by changes to national science curricula in the UK that affected test-takers’ interpretation of the test specification. The assessment context for BMAT is described to demonstrate how a review can facilitate positive impact by design (Saville, 2012), and the steps taken to align the revised specification with best practice in supporting learning through positive washback are outlined. The process was informed by surveys of BMAT candidates on preparation behaviours, and findings from these are briefly presented to illustrate the rationale for aspects of the revision. We reference in particular international candidates who may take BMAT for entry to biomedical courses in their home country or overseas. Also discussed are considerations for candidates whose first language is not English, and the challenges of ensuring equitable assessment of test-takers who have experienced pedagogic variations in science teaching and conceptualisations of scientific knowledge. Furthermore, we present example issues identified during the review, to highlight the value of engaging in this resource-intensive process to test providers and educational policymakers.

**Validity in admissions testing**

Research on admissions testing tends to focus on criterion-related validity, particularly on predictive validity using academic course performance as outcome criteria (Soares, 2012). As admissions tests aim to select those with the best chances of succeeding academically, this is an understandably important area. In the US, over a century of admissions testing has shaped the development and revision of large-scale tests. Despite continuing debates surrounding their use and validity, there is now some consensus that predictive validity should be supplemented with face validity (Atkinson & Geiser, 2009; Linn, 2009).

Compared to the US setting, admissions testing in the UK does not have as long a history nor has it been as widespread. Medical selection is one context where testing
became more widely accepted before other subjects, as some UK medical schools have used BMAT or its predecessor, the Medical and Veterinary Admissions Test (MVAT), since 2000 to aid selection of students (Emery & Bell, 2009). The majority of research on UK medical selection testing has concentrated on predictive validity (Emery & Bell, 2009; Kelly et al., 2013; McManus, Dewberry, Nicholson, & Dowell, 2013). However, studies have shown that other aspects of validity should not be overlooked in medical admissions, as the skills assessed by some tests are not seen as important by applicants and members of the medical profession, indicating low face validity (Cleland, French, & Johnston, 2011; Kelly, Gallagher, Dunne, & Murphy, 2014; Stevens et al., 2014). The negative perceptions identified in these studies focused on abstract reasoning sections of assessments, which were viewed as having little relevance to studying medicine. Therefore, it is crucial to consider the perceived relevance of constructs tested for medical selection, as they are abilities developed by candidates when they prepare for tests.

Further, Messick has argued that any model of test validity that did not account for the wider impact of the test on candidates and society was inadequate, as it failed to account for ‘both evidence of the value implications of score meaning as a basis for action and the social consequences of score use’ (Messick, 1995, p. 741). This observation applies to medical admissions tests, and validity arguments need to account for the social consequences of using admissions tests in selection, including the ways that candidates prepare. Test providers have a responsibility to examine aspects of validity other than criterion-related components alone. This can be facilitated by adopting and tailoring frameworks of validity established in other areas of educational assessment.

A framework for the validity of admissions tests

Validity is a multi-faceted construct, and establishing the validity of any test requires many sources of evidence of its use to be evaluated as a whole. The merits of various validity frameworks are much debated and contested (Newton & Shaw, 2014). Frameworks by Kane (2006; cited in Cook, Brydges, Ginsburg, & Hatala, 2015) and Messick (1989a) are frequently applied in medical education and emphasise that validity is unitary as an ‘integrated value judgement’ (Messick, 1989b, p. 5) on the appropriateness of an assessment. The field of language assessment has led the development of socio-cognitive approaches to validity which address the mental processes that candidates engage in, whilst also recognising that test items must be situated in, and reflective of, the real-world contexts in which candidates apply their skills. Weir’s (2005, 2013) model of test validity has been used extensively by Cambridge Assessment English as the framework for establishing the validity of English language assessments that they develop (Weir, 2013). As part of the Cambridge Assessment Group, Cambridge Assessment Admissions Testing has also adopted this framework as a way to conceptualise validity in tests for university selection, due to the comprehensive and structured treatment it provides of multiple aspects of the testing process (Cheung, McElwee, & Emery, 2017).

Similar to Kane and Messick’s frameworks, Weir’s framework also emphasises the importance of integrating multiple sources of validity evidence and provides a
structured method of assembling and evaluating this evidence. This includes issues pertaining to the physical, psychological and experiential characteristics of candidates, the representativeness of the test tasks with reference to the intended construct, characteristics and administration of the test tasks, and the appropriateness and reliability of the scoring. The model also covers external evidence that the test relates to the construct of interest, and the impact of the test and test scores on various stakeholders, including test-takers. This final facet of validity is of particular interest in this paper. Figure 1 illustrates the framework, where ‘the arrows indicate the principal direction(s) of . . . hypothesised relationships . . . and the timeline [of the various phases of establishing validity]’ (Weir, 2005, p. 43). The implied sequential nature reflects the reality for test developers that certain elements of validity evidence cannot be collected until after a test event.

**Consequential validity**

Weir’s (2005, 2013) socio-cognitive model includes consequential validity as a crucial piece of evidence for scrutinising the fitness for purpose of any test and is influenced by Messick’s (1989b) concern with the consequences of test use (Hawkey, 2011). Consequential validity is typically divided into two elements: washback and impact.

Impact is the effect that the test has on the full range of stakeholders and on society more generally. Test-takers and selecting institutions are those affected most directly, as the results influence decisions about their future study paths and careers, and their academic cohorts, respectively. Schools, parents, and national medical and dental associations represent other groups impacted by tests such as BMAT in the wider social sphere.

Washback refers to the extent to which an examination influences learners’ preparation materials (Saville & Hawkey, 2004), the teaching they receive (Green, 2007b), the
curriculum they follow (Cheng, 2005; Qi, 2007), and their perceptions of exams (Gu & Saville, 2016). Assessment can also impact educational policy (Shohamy, 2007) and student mobility around the world (Yu & Jin, 2016). Washback has directionality (Alderson & Wall, 1993) and can be positive, encouraging candidates to study widely and promoting the acquisition of beneficial skills and knowledge, or negative, directing the student to concentrate on narrow aspects of the curriculum, to attempt to ‘question spot’, or to focus on test strategies more than on learning. While test scores are assumed to represent ability in a certain domain, where preparation effort is focussed on specific elements due to the expectation that certain content will be examined, it may not be justifiable to infer mastery of the broader domain.

Hughes (2003) suggests that positive washback in high-stakes tests can be achieved by testing abilities whose development you want to encourage, by sampling widely from the curriculum and by ensuring that the test is known and understood by students and their teachers. Therefore, a test’s specification of curriculum and revisions to it are a direct way that test designers influence test-takers’ preparations.

One interesting difference between washback in language testing and washback in admissions testing is the role of teacher support and provision in test preparation. Green (2007a) and Hughes (2003) refer to the importance of teachers being familiar with the test content. This means that studies of language testing impact and washback often concentrate on school settings and the efforts of teachers to prepare their students for assessment (e.g. Saville & Hawkey, 2004). By contrast, admissions tests such as BMAT are recognised as sitting outside of formal school examinations. A key concern of universities that use BMAT is that preparation should not entail significant new learning. Excessive time spent on test preparation could detract from developing other competencies relevant for biomedical study.

As relatively few students sit admissions tests, the resources available to help students prepare can be limited. However, there is a perception that students in fee-paying schools are more likely to receive preparation help (Cleland et al., 2011). To counter this, it is crucial to make supporting materials freely and widely available to all candidates, so that test-takers from schools entering candidates regularly do not have an advantage through more familiarity with test materials. The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) references the accessibility of information (Standard 8.1) as central to fairness for candidates, stating: ‘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). Cambridge Assessment Admissions Testing makes all of its test preparation materials available free of charge on its website. The changes to BMAT Section 2 and the development of supplementary revision materials were intended to further ensure a more level playing field for candidates, and to act as a substitute for teacher intervention.

A challenge for the test provider is how to maximise positive washback where the teacher’s role is diminished, and candidates mostly engage in test preparation outside of school. In language learning, test washback is sometimes conceptualised as a lever for positive educational change that can influence curricula in language schools (Andrews, 2004). For subject-specific admissions tests, washback effects are more likely to be
concentrated directly on candidates; however, school-taught curricula must still be considered carefully by the test provider, so that time spent on test preparation is beneficial for the test-taker in other examinations.

The steps taken to embed good practice in terms of positive impact and positively influencing stakeholder acceptability are outlined. Sharing the review demonstrates how the concepts of test impact and washback can be considered in the admissions testing context, where they are less frequently discussed than in the language testing domain.

**Context and assessment focus of BMAT**

BMAT is used by 34 UK and international universities to help select candidates into over 45 medical, biomedical science, and dental programmes. Candidates’ BMAT scores are typically considered alongside performance on a range of other measures such as school grades, personal statements of suitability for biomedical study, interviews, or multiple mini interviews.

As an admissions test BMAT has four aims:

- to differentiate between large numbers of applicants who attain highly in school examinations;
- to ensure that applicants’ scientific understanding is adequate to cope with the initial years of rigorous science-based study in their biomedical course;
- to provide a common measure to compare applicants from a wide range of educational backgrounds and with a variety of qualifications, including overseas applicants, mature applicants, and applicants from different school types, many of whom only have predicted grades at the point of application;
- to support focus on applicants who have a realistic chance of receiving an offer on their chosen biomedical course and obtaining the final school grades to take up that offer.

*Figure 2* provides an overview of BMAT’s sections, format and timings.

**Preparation for BMAT**

There is often a lack of clarity in the medical education literature around terms such as ability, achievement and aptitude. While there is frequently ambiguity in their use, they have specific meanings. Ability refers to the current level of performance. Achievement refers to demonstrated prior attainment. Aptitude is the potential for developing a skill or competency and has historically been associated with innate characteristics (Cheung & McElwee, 2017; Newton & Shaw, 2014). Such distinctions are important as test-takers may infer messages about the test construct and the value of preparation according to their use. Fisher (2005) outlined the Cambridge Assessment approach to admissions testing when describing the Thinking Skills Assessment, an admissions test whose construct is similar to BMAT Section 1: ‘the skills assessed ... are teachable skills which are valuable in higher education ... [L]earning thinking skills leads to the enhancement of students’ performance in all parts of the curriculum as well as preparing them to get the best from Higher Education’ (2005, p. 12).
Despite the legacy term ‘Aptitude and Skills’ in the title of BMAT Section 1, BMAT’s construct is not based on innate skills and it is not a general intelligence test; the content in both Sections 1 and 2 is directly related to skills valued in school academic performance and therefore preparation behaviours should map to skills of wider educational benefit for the candidate. While preparation for the test should comprise mostly revision and familiarisation with the test format, it is important to acknowledge how BMAT preparation has changed since the introduction of the test and how this may impact test-takers, particularly those from overseas.

Cambridge Assessment Admissions Testing has conducted surveys at various points to inform the development of preparation materials, and to monitor trends in how students prepare (McElwee, Fyfe, & Grant, 2017). A 2008 survey of successful medical school applicants indicated that candidates’ median self-study time for BMAT was approximately 8 h, with 50% of participants spending between 5 and 15 h preparing (Emery, 2010). A survey conducted through the BMAT website immediately after the 2015 test suggested that the median self-study time had increased to 30 h (Gallacher, McElwee, & Cheung, 2017). The surveys are not directly comparable in terms of sample and the point in the admissions process at which data were collected, and issues such as the length of time between the test and survey in 2008 may have affected estimates. However, the increase in preparation time is notable.

A minority of candidates reported getting some preparation help from their school, although this usually amounted to less than 3 h, supporting the idea that washback effects from BMAT primarily impact learners’ self-preparation and revision. In both the 2008 and 2015 samples, however, test-takers from UK independent schools (fee-paying) were more likely to receive preparation help.

The issue of test preparation gained increasing importance during this time for a number of reasons. Renewed efforts to increase diversity and widen participation in the UK medical profession (Medical Schools Council, 2014; Milburn, 2012) emphasised the importance for test providers to investigate fairness and bias with research (e.g.}

<table>
<thead>
<tr>
<th>Section 1</th>
<th>Tests generic skills often utilised in undergraduate study: problem-solving, understanding argument, and data analysis and inference skills. There are 35 multiple-choice questions in 60 minutes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 2</td>
<td>Tests whether candidates have the core knowledge and the capacity to apply it, which is a pre-requisite for high level study in biomedical sciences. Questions are restricted to material typically included in non-specialist school Science and Mathematics courses up to the age of 16, but will require a level of understanding appropriate for such an able target group. There are 27 multiple-choice questions in 30 minutes.</td>
</tr>
<tr>
<td>Section 3</td>
<td>Tests the ability to select, develop and organise ideas and to communicate them in writing, concisely and effectively. A selection of questions on topics of general, medical, or scientific interest will be available, one of which must be chosen. The response is limited to one A4 page in 30 minutes.</td>
</tr>
</tbody>
</table>

Figure 2. Overview of BMAT sections, question formats and timings.
Emery, Bell, & Vidal Rodeiro, 2011; Lambe, Waters, & Bristow, 2012; Tiffin, Dowell, & McLachlan, 2012), to ensure that selection tests did not disadvantage particular groups, including in terms of test access and preparation. The characteristics of the test-taker population of BMAT also changed between 2008 and 2015, as a growing number of institutions in the UK and internationally began using BMAT to support their selection processes, increasing the proportions of candidates who had completed their schooling outside of the UK and who spoke English as a second or additional language.

The issues around preparation, widening participation and international applicants informed the review and revision of BMAT Section 2 throughout. However, the impetus for the review was also influenced by changes to UK science curricula.

**Rationale for a new BMAT specification**

BMAT was used initially only by universities in England for selection and therefore the National Curriculum for England, Wales and Northern Ireland, which outlined the compulsory science curriculum for state schools up to age 16, was an appropriate basis for the content of the Scientific Knowledge and Applications section. This curriculum content was reflected in the GCSE\(^1\) specifications from UK examination boards and accompanying textbooks. The information on the examinable topics in the BMAT specification was deliberately brief, directing candidates to the official National Curriculum documentation. The intended message to test-takers was that the core knowledge required for BMAT was already familiar to them through their GCSE preparation or other schooling up to the age of 16, and therefore it was not necessary to learn significant amounts of new material; preparation should comprise revision of previously learned topics. However, later changes to the National Curriculum resulted in a less prescriptive programme of study up to the end of compulsory schooling and national GCSE examinations. The five GCSE examination boards (OCR, Edexcel, AQA (England), WJEC (Wales) and CCEA (Northern Ireland)) had increasing freedom to specify the content and curricular routes to achieving GCSEs in Science and Mathematics. This posed a challenge for constructing BMAT Section 2 to ensure that the questions contained content covered in most UK state schools by age 16, particularly in Biology where BMAT Assessment Managers found there was most divergence between the examination boards’ syllabus specifications.

The less detailed National Curriculum documentation was a threat to the validity of BMAT. Candidates would have had more difficulty identifying topics that might be included in the test for preparation, violating Hughes’ (2003) observation that positive washback is encouraged when the test demands are known and understood by students. This issue was compounded for non-UK candidates following other international school curricula, for whom the challenge of understanding which topics might be covered in BMAT Section 2 was arguably more difficult. More importantly, from a pedagogical perspective, it was recognised that international science teaching is not uniform – different problem solving approaches, assessment focusses, linguistic structures, and students’ assumed knowledge impact how candidates interpret questions, affecting their difficulty (e.g. El Masri, Baird, & Graesser, 2016).

Cambridge Assessment decided that a full review of BMAT’s Section 2 assessment content and the development of a more detailed specification were necessary to ensure
that test-takers were fully supported in their preparation for the test. This exercise also presented the opportunity for consultation with the universities using BMAT to ensure that the topics represented were those most relevant for biomedical study. This maximised stakeholder acceptability for both the using institutions and test candidates.

The review and revision process

The review of BMAT Section 2 consisted of six main phases of work (Figure 3), which are described here to provide a methodological framework for similar projects. They illustrate examples of how the test specification was designed to best support face validity and consequential validity through consideration of candidate preparation, alignment between school and biomedical course curricula, and a focus on developing candidates’ understanding of the breadth of the discipline and the relationships between various scientific topics. The revised specification was first used for the November 2014 BMAT session.

Identification of relevant GCSE specifications from UK examination boards

Prior to the review, Section 2 items on BMAT aimed to use topics and knowledge that were likely to have been encountered by candidates who took UK GCSEs as the basis for the problem solving elements of their questions. Thus, the first step was to identify a sample of the examination boards’ science curricula from which to derive the BMAT specification. GCSE Mathematics, Science Double Award and Triple Award\textsuperscript{2} pathway specifications from each UK board were chosen.

Review by BMAT senior examiners

Each subject area in BMAT Section 2 (Mathematics, Biology, Physics and Chemistry) has a team of subject experts who have experience of question setting and marking UK and international GCSE qualifications or undergraduate teaching. A senior member of each team reviewed the examination boards’ specifications in their subject area to identify commonalities and establish the core content of the draft specification. The

![Figure 3. Outline of the review and revisions process.](image-url)
general topic area, details of sub-topics and extent and nature of the commonalities were documented (e.g. depth of knowledge expected on the topic, diversity of exemplars used, and which boards did not include the topic). A principle was followed such that all topics included in BMAT should be represented in three (preferably four) of the five awarding bodies' syllabuses. Additional topics were recommended for inclusion that did not have significant overlap between the specifications or did not appear on any GCSE specification, where the examiners judged that:

- a topic was essential to a core understanding of the particular science;
- a topic was fundamental to the relevant scientific principles, and taught earlier than GCSE; such topics should be explicitly included for students' reference, and particularly to support international candidates who may encounter certain principles taught in different ways;
- examples drew links between topics to promote understanding of the interrelatedness of science as a discipline.

To warrant inclusion, the topic should be encountered by students by the end of the school year immediately preceding BMAT, or it should be judged to be accessible and manageable for able science students to learn independently using a revision guide. Ideally, these topics were also of particular relevance to medical or biomedical sciences.

BMAT subject assessment managers compiled and reviewed the first draft of the specification to ensure no duplication of topics or obvious deviations from the agreed parameters of the review. In addition, the review considered subject curricula from equivalent Scottish Qualifications Authority (SQA) examinations, to ensure that the selected topics were suitable for those completing their schooling in Scotland.

**Consultation with biomedical lecturers and admissions tutors**

The draft BMAT specification was sent for consultation to representatives of the universities that use BMAT. The content was further refined and agreed during a full day round-table discussion bringing together senior academics involved in the selection and education of undergraduates in medicine or related biomedical sciences at these universities. Thirty-nine broad topic areas had been distilled from the examiners' initial review, comprising 507 sub-topics or pieces of information. Three questions were used to guide the discussion about each topic:

- Are the knowledge and concepts important for medical/biomedical study?
- Is the context provided for the underlying principles relevant?
- Is the type of thinking that a topic affords important or useful in studying medical or biomedical subjects?

The first two questions capture issues relating to the construct validity of BMAT Section 2 and ensuring that its curriculum topics are relevant to the study of medicine or biomedical sciences. Amendments to the draft specification ensured these principles were embedded. For example, in Physics, the topic of electricity was common across all exam board syllabuses. The academic panel agreed its retention on the BMAT
specification, for its relevance to biomedical topics such as nerve impulses. However, the sub-topic of *domestic electricity*, which did not have that same relevance, was excluded. A second example is the topic of cosmology, which was removed from the initial draft specification for Physics in BMAT. However, key ideas associated with the topic such as the Doppler effect and line absorption spectra were identified as particularly relevant to biomedical study and were re-categorised under the topic areas of ‘Waves and Wave Behaviour’ and ‘the Electromagnetic Spectrum’, respectively. These modifications were intended to ensure revision time is not focused on topics that will not be relevant to later biomedical study. As BMAT tests students’ understanding of scientific principles rather than specific facts, this approach also encourages abstraction and generalisation of principles beyond the original context in which they were learned, reflecting real-world applications and encouraging advanced scientific thinking.

A further objective of this stage in the review was to identify examples given for core topics by exam boards. Each GCSE examination board included homeostasis in their Biology curriculum but the illustrative examples were different in each case (e.g. a combination of regulation of blood glucose, temperature, or water content). This meant that, prior to the introduction of the new specification, BMAT questions on this topic could not assume knowledge of any particular homeostatic system and thus could be focussed only on the general principles of homeostasis. The revised BMAT specification explicitly stated the regulation of blood glucose, water content and temperature as the examples on which questions could be based.

The opportunity to review the draft specification alongside older National Curriculum documentation led to the academic panel’s recommendation to include organs and organ systems in the biology specification and to differentiate between the biochemical aspects of respiration and the physical (mechanical) actions of breathing including the thorax and lungs. Following review and discussion with the academic panel, the draft specification was refined to 35 topics and 196 subtopics.

**Review by international curriculum experts**

Specifications and materials from Cambridge Assessment International Education, a part of the Cambridge Assessment Group that develops international school science curricula, were obtained for review. Science and Mathematics education experts from countries with significant interest in BMAT were also consulted to check the appropriateness of the specification and suggest any adjustments that might be needed. As BMAT was used by the University of Leiden in the Netherlands and has since been adopted by other Dutch universities for entry to biomedical science courses, the consultation with Dutch educators was particularly important. A panel of secondary school science teachers, some of whom also worked as university tutors, were selected for their familiarity with both the curricular routes that students might take in science learning at school, and the expectations of entrants to medicine in terms of their science knowledge. They reviewed the draft specification and through discussion with BMAT Assessment Managers advised on issues that could affect the interpretation or difficulty of questions for candidates. These differences were distinguished as those driven by linguistic issues, those reflecting differences in science teaching approaches, or a combination of the two. Examples of each from the Dutch context will be outlined.
**Difference in teaching approaches**

One example demonstrating the importance of understanding teaching approaches comes from Mathematics. In the UK, students typically learn to solve quadratic equations using factorisation, whereas in the Netherlands the usual approach is through the use of the quadratic formula and substitution of values into the equation. This means questions containing numbers that are easily factorised, a relatively trivial step for UK students, could be significantly harder for students attempting the substitution method.

A further difference is that Dutch students are permitted to consult a reference handbook in examinations and therefore tend not to commit key facts and formulae to memory, as would be expected of UK students. BMAT does not allow use of additional reference guides, so supplementary material was compiled for students taking BMAT for entry to Dutch universities to support their preparation, advising of differences in curricular approach to certain topics, and highlighting areas of the BMAT specification where key formulae should be memorised.

**Issues with linguistic elements**

The review identified some linguistic issues related to the use of scientific terminology in the Netherlands. One important consideration is that the universities in the Netherlands that use BMAT teach their courses in Dutch so candidates are not necessarily preparing for BMAT with a view to studying through the medium of English. Even though most Dutch candidates are highly proficient speakers of English, some scientific and mathematical vocabulary remain that they might not have been reasonably expected to have encountered before. Therefore, experts familiar with the Dutch science curriculum assess BMAT Section 2 questions for scientific vocabulary that candidates would not be readily able to translate; these terms are provided in a short glossary for BMAT candidates in the Netherlands who are applying to Dutch universities.

**Issues stemming from the interaction of teaching approaches and linguistic elements**

In some cases, international curriculum experts identified issues resulting from an interaction of linguistic expression and teaching approaches. This is exemplified in the common Chemistry topic of oxidation and reduction, which concerns how a specified chemical species oxidises other atoms/compounds. It does this by taking electrons from those atoms and the specified species itself is reduced in the process, so something that is an oxidising agent is itself reduced and receives electrons. Equally, something that is a reducing agent loses electrons and is itself oxidised.

In the UK education system, students are taught to remember this by learning about what is happening to the other atoms/compounds. However, in the Dutch language, the focus is on what is happening to the specified chemical species. This impacts the way that an item written in English is interpreted by a Dutch student. The statement 'the metal oxidises' is easily understood by an English student as the metal acting upon something else – in other words, 'the metal oxidises some other atom/compound'. On the other hand, a Dutch student, even one with a high level of English language proficiency, is likely to interpret the same statement as 'the metal itself was oxidised'.
Disentangling the construction of these statements can induce a higher cognitive load, which would increase the difficulty of an item artificially. Identifying this issue allowed Assessment Managers to make small changes to phrasing for items on this topic and avoid any potential confusion.

**Reviewing the language demands of BMAT for international candidates**

In other international contexts, such as Singapore, BMAT tends to be used for courses where English is a medium of instruction. BMAT candidates are expected to have a level of English language proficiency commensurate with the requirements for studying biomedical courses in English. While the questions in all three sections of the paper reflect this, it is important that the language is not so complex as to introduce construct-irrelevant variance by discriminating against candidates who are not native speakers. Non-native speakers of English taking courses where English is the medium of instruction are typically required to evidence an IELTS score of 6.5. This is roughly equivalent to Level B2 on the Common European Framework of Reference for Languages (CEFR) and is the specified level for the English proficiency of BMAT candidates.

The subject-specific science context of the questions in Section 2 of BMAT means that question stems tend to be short and quite straightforward. In BMAT Section 1, questions relating to critical thinking and understanding argument have longer and more complex texts. Assessment Managers with expertise in English language assessment (as part of the wider Cambridge Assessment group) train item writers and vetters on the differences between B1, B2 and C1 on the CEFR for reading comprehension. At pre-editing, editing and vetting stages, Assessment Managers, item writers and vetters check BMAT Section 1 to ensure that the text and grammatical structures are consistent with a degree of difficulty no higher than Level B2 on the CEFR. They also evaluate each item specifically to check that difficulty does not come from the way the item is expressed, the reading load or cultural assumptions that may be unfamiliar to international candidates.

In BMAT Section 2, where an Assessment Manager or editing panel considers the language level of an item is too high (particularly where the test is intended for use in selection to courses where English is not the medium of instruction, such as The Netherlands), a range of adjustments are employed to ensure that items are suitable. These include:

- providing glosses and paraphrases of unfamiliar terms, if they cannot be avoided;
- making the register less formal;
- reducing sentence length and simplifying the syntax of a passage or input;
- reducing the density of information in a passage or input;
- replacing UK-specific names, institutions or customs with generic equivalents.

**Specification trialled by item writers**

In the final stage of specification development, BMAT item writers trialled the draft specification and were asked to assess whether the breadth allowed them to write high-quality questions that tapped scientific problem solving in novel ways not typically encountered in school examinations. As part of the iterative process of compiling the specification, they were asked to comment on any elements they felt were missing that
would help to link topics or significantly improve the quality of questions that could be generated. The feedback from the item writers was favourable. Item writers for Biology were particularly positive, as the lack of overlap of exemplars between examination boards’ curricula had previously limited the set of topics on which questions could be set.

**Development of free revision guide**

To supplement the specification development, Cambridge Assessment Admissions Testing also developed a free revision guide for candidates that was published as the assumed subject knowledge guide. Coordination Group Publications (CGP), a well-known publisher of GCSE and A-level revision guides, was selected as a partner in this exercise. CGP guides are highly visual rather than text driven, and would likely be recognisable to many UK students, reinforcing the idea that preparing for BMAT should be revision rather than new learning.

The resulting *Assumed Subject Knowledge Guide* for BMAT Section 2 had implications for positive impact for the UK and international candidates by attempting to reduce the advantage for test-takers who receive school preparation support for BMAT. For international candidates, the revision illustrates the way that critical topics are assumed to have been taught and provides a guide to key scientific terminology that candidates could encounter, both in BMAT and later on in their biomedical degree. For those candidates who do not speak English as a first language, this specialised vocabulary resource is an important means of accessing both test preparation and the test questions themselves. The original revision guide was available as an e-book free of charge to any prospective BMAT candidates who registered to receive it. A newer version, produced wholly by Cambridge Assessment Admissions Testing has recently been launched on the Moodle platform and offers improved accessibility, now allowing candidates to download and print sections for further study. Registration for the test itself is not a condition of access, so test-takers can revise from the material online before they make a final decision about taking BMAT.

**Conclusion and recommendations**

The revision process drew from a validity framework widely used in language testing (Weir, 2005) to include socio-cognitive aspects of validity in the evaluation of an admissions test, enabling BMAT developers to identify issues that could otherwise have been overlooked. The consultative approach taken with academics in medical and biomedical disciplines ensured face validity of BMAT, both for using institutions that can see the direct link to knowledge and skills needed for successful progress in medical degrees and for candidates who invest time in preparing for the test. This process was intended to support cognitive components of validity by attempting to align the reasoning required in BMAT to that required in studying for biomedical degrees.

Approaches to validity that focus on social aspects of the test, such as consequential validity (Green, 2007a), informed the review and the design of the specification. Starting with common GCSE science and maths curricula was intended to reduce the amount of
additional learning that preparation for BMAT might need, and to make BMAT revision useful for other exams, as well as future biomedical study. The principles outlined by Hughes (2003) to support positive washback were incorporated in the process, with consideration given to how principles valued by stakeholders, such as an understanding of how science topics inter-relate across subject areas, could be reflected in the new BMAT curriculum. Additionally, development of the assumed subject knowledge guide is an example of impact by design (Saville, 2012), as its publication was intended to make detailed preparation guidance materials available for all BMAT candidates.

Access to professions such as medicine is often considered opaque and the multi-stage selection process has been identified as a barrier to applicants from certain demographics. For international candidates, particularly those who do not speak English as a first language, preparation for admissions tests, as one part of this process, can be a particular challenge. The review of the BMAT Section 2 curriculum highlighted the particular linguistic challenges that speakers of other languages may face in preparing to take BMAT – specifically around the construction and complexity of language used in the test items and the occurrence of subject-specific vocabulary that may be significantly different to that used in the candidate’s L1. Further, we identified a more complex interaction between linguistic and pedagogical issues that could affect candidates’ conceptualisation of certain science topics and calculations, differentially impacting the difficulty of questions for speakers of languages other than English. Insight into these issues allowed Assessment Managers and item writers to take specific steps to improve the accessibility and experience for international candidates preparing for this admissions test in English.

Limitations of the review

The revision of the BMAT Section 2 syllabus was intended to increase transparency, and support international candidate preparation and widening participation. However, the extent of its success in fulfilling these aims is currently unknown and it is important to acknowledge the limitations of this piece of applied work, particularly to inform future planning and similar projects.

The initial surveys examining preparation behaviours for BMAT, which provided part of the rationale for the revision, have some limitations. The first survey was carried out with only successful applicants to medical school, some time after they had taken BMAT, factors which may have affected respondents’ recollections, while the second survey was a convenience sample, not necessarily fully representative of the full BMAT cohort. Therefore, the results of these surveys have been interpreted cautiously. Registrations for the Assumed Subject Knowledge revision guide suggest that the majority of BMAT candidates access this free resource, and candidate reactions have been broadly positive but candidates’ use of the guide has not been universal. Different technology systems underpin the registration and revision guide delivery mechanisms and so revision guide use may be underestimated. There may also be issues around raising awareness of the availability of the resource, its usability, or candidates’ access to technology – the original version of the guide was available only as an e-book, whereas a relaunched edition now has sections that can be downloaded, saved and printed.
Further work is needed to investigate the demographic variables associated with candidates’ use of the guide, and to explore BMAT preparation behaviours in more detail. Finally, the complexity of the evolving widening participation agenda in higher education in the UK makes isolating the impact of this revision to assess its contribution difficult. Although the wider impact needs further evaluation, some benefits from the review are immediately apparent.

The examples presented demonstrate the value of reviewing admissions tests with consideration of linguistic issues, even when assessing candidates who are fluent in the language of the test, and the teaching practices of candidates’ home countries. These factors, as well as acknowledgement of the complexity of the interaction between pedagogic and linguistic issues, had a substantial impact on the review. Based on the range and complexity of issues identified, we recommend that providers of admissions tests that have sizeable international candidatures consider the language demands and linguistic issues in reviews of their tests. They should also attempt to mitigate the potential impact of international differences in teaching practices by using local subject experts, particularly when testing mathematical, scientific or quantitative reasoning. Given the importance of widening participation in the study of medicine (Medical Schools Council, 2014), test developers have a responsibility to support candidate preparation and ensure equity of access to supporting materials. As higher education becomes increasingly internationalised, with greater student mobility, the numbers of candidates preparing for admissions tests in languages other than their first language, and most frequently tests in English, will grow and the specific challenges this raises in terms of fairness and validity must be addressed by testing organisations. Furthermore, we encourage the use of broader approaches to validity in admissions testing. The review presented in this paper demonstrates that established frameworks from language testing can support these aims by ensuring that test providers consider consequential validity, washback and impact in the design and administration of their tests.

Notes

1. General Certificate of Secondary Education; the public examinations taken by most UK students at approximately age 16.
2. UK GCSE students were entered for Science Single Award, Double Award or Triple Award, where they received one, two or three GCSEs for their combined study of Biology, Physics and Chemistry. Science Double Award was the variant for which UK students were most commonly entered.
3. IELTS: International English Language Testing System. IELTS is a high-stakes language test used for study, migration and work. Language proficiency is measured on a nine-band scale.

Disclosure statement

The authors are employed by, or have previously been employed by, Cambridge Assessment which owns and administers the BioMedical Admissions Test (BMAT).
Notes on contributors

Sarah McElwee is a Principal Research Manager at Cambridge Assessment Admissions Testing and is responsible for managing the research team working on BMAT and other admissions tests to higher education. She holds a PhD in Cognitive Development from Queen’s University Belfast. Her current research interests include the consequences of high-stakes testing and the development and assessment of thinking skills.

Kevin Y.F. Cheung is a Principal Research Manager at Cambridge Assessment Admissions Testing and a Chartered Psychologist. He holds a PhD from University of Derby on the role of identity in academic writing. His research interests include construct validity in admissions testing and the role of standardised testing in multi-method selection processes.

Stephen R.T. Cromie is a Senior Assessment Advisor at Cambridge Assessment International Examinations. He previously worked as a Senior Assessment Manager for Science in Cambridge Assessment Admissions Testing. Stephen holds a PhD in Theoretical Physics from Queen’s University Belfast.

Mark Shannon is Group Assessment Manager for Sciences at Cambridge Assessment Admissions Testing. He manages the team responsible for the development and construction of a range of science and mathematics admissions tests, including BMAT Section 2.

Tom Gallacher worked as a Research Analyst at Cambridge Assessment. He was responsible for statistical analysis that supports research and validation activities for admissions tests including BMAT. He holds an MSc in Organisational Psychology from the Manchester Business School and an MA in Psychology from Edinburgh University.

ORCID

Sarah McElwee http://orcid.org/0000-0001-7734-8249
Kevin Y.F. Cheung http://orcid.org/0000-0002-9548-2932
Tom Gallacher http://orcid.org/0000-0002-7556-5327

References


This document was initially designed for print and as such does not reach accessibility standard WCAG 2.1 in a number of ways including missing text alternatives and missing document structure.

If you need this document in a different format please email admissionstesting@cambridgeassessment.org.uk telling us your name, email address and requirements and we will respond within 15 working days.