The best choice?

Katherine Woolf, Henry Potts, Josh Stott, Chris McManus, Amanda Williams and Katrina Scior consider evidence on selection into the healthcare professions

How do we select people to become healthcare professionals? Psychologists, medical doctors and other healthcare professionals have demanding jobs that significantly impact on the public. The professions, and academic psychologists, have a duty to produce evidence on who should fill these roles. Many of us are personally involved in choosing candidates and want to know if our choices are sound. It is now widely accepted that those choices must be carefully balanced with the mandate to increase demographic diversity (e.g. Medical Schools Council, 2014).

In many professional courses, such as clinical psychology and medicine, the validity of selection methods is especially important because nearly everyone accepted qualifies to practise. Training routes are heavily oversubscribed, with the applicant-to-place ratio for clinical psychology as high as 30:1. This has led to concerns that it seems next to impossible to gain a place, and that selection processes are opaque.

This article provides an overview of selection methods, asking how well they predict later performance and how fair they are. We use examples from medicine, where there is considerable research evidence (cf. McManus, 2005), and from clinical psychology, where there is less, but which is more relevant. Medicine and clinical psychology are interesting because they require selectors to consider candidates’ personal qualities such as empathy, warmth and a commitment to best practice, as well as academic ability.

Assessing candidates

The UK’s 30 clinical psychology and 33 undergraduate medical programmes use various selection methods. Academic ability is evaluated using GCSEs and A-levels or equivalents in medicine – a minimum of three As (and increasingly at least one A*) at A-level and a B at AS-level being required. Degree performance is more important in clinical psychology, a good 2:1 or first class degree being required and postgraduate qualifications providing an advantage. Interviews, personal statements and referee ratings are used to assess person-based qualities. Recently, situational judgement tests – multiple-choice tests that ask candidates how they would behave in relevant situations – have been used to select medical students for junior doctor jobs, and there is pressure for them to be used in clinical psychology selection (Health Education England, 2014). Medical schools also use aptitude tests such as the mUltiple mini-interview in the UK to predict future performance in practice.

How can we tell which selection methods work best for healthcare education?

How can we predict whether someone will become a good healthcare professional?


A key way to test whether selection methods work is to see whether people with high scores at selection subsequently do well, and those with low scores do badly.
UKCAT (UK Clinical Aptitude Test) and BMAT (Biomedical Admissions Test) to measure academic and personal qualities because so many applicants have top A-level grades and because they are, it is claimed, fairer.

A key way to test whether selection methods work is to see whether people with high scores at selection subsequently do well, and those with low scores do badly. This provides evidence of predictive validity. For example, if students with three A* grades at A-level do better on course assessments than those with three A grades, and the latter do better than those with three B grades, this helps justify rejecting candidates with lower grades who could reasonably be expected to fail the course.

Of course it is not that simple, because as well as using diverse selection methods, professional courses teach and assess their students’ academic and person-based knowledge and skills in various ways. Clinical psychology is a three-year postgraduate full-time course providing academic teaching and study, clinical placements and research training. Assessments include reports on clinical work, placement supervisor ratings, a doctoral thesis viva, exams, essays and clinical simulations. Medicine is usually a five-year undergraduate course. Students learn the scientific underpinning of clinical medicine in a university setting, and learn practical skills and professional behaviour in NHS or community settings. Assessments are typically multiple-choice written examinations and practical face-to-face examinations. In choosing a selection toolkit, we need to know which selection methods predict academic outcomes, which predict person-based outcomes, and which selection methods add value to others and should be used in combination.

Selecting on academic ability
A-level grades are good predictors of undergraduate performance across undergraduate subjects (HEFCE, 2014) and in medicine specifically (Ferguson et al., 2002; James & Chilvers, 2001; McManus, Woolf et al., 2013; Yates & James, 2007). A-levels are (usually) written knowledge tests and thus predict performance on in-course written knowledge tests well, although they also predict performance on practical tests of professional skills and behaviours. A study of 700 medical students found high A-level grades (along with high conscientiousness and white ethnicity) were better predictors of high scores in written and clinical examinations than study habits, parental socio-economic status, speaking English as a first language, or sex (Woolf et al., 2012). Data from 12 medical schools showed that up to 65% of the variance in first-year scores (mostly written knowledge tests) was accounted for by A-level performance (McManus, Dewberry et al., 2013a); and a meta-regression of six studies of medical school entrants estimated the predictive validity of A-levels to be very high, around 0.8 (McManus, Dewberry et al., 2013b). Data from seven cohorts of the University College London DClinPsy course showed that better A-levels predicted higher marks on all four in-course exams, and were associated with fewer reported concerns about clinical placement performance (Scior et al. 2013). GCSEs have less predictive validity (McManus, Woolf et al., 2013), although when nearly all applicants have top A-level grades they add useful information (Woolf et al., 2012).

Theoretically, A-levels and GCSEs are good predictors of academic and person-based performance in medicine because they provide the foundation of what has been called the ‘academic backbone’: the accumulation of...”medical capital”; that set of knowledge, theories, experience, understanding and skills that comprise successful medical practice’ (McManus, Woolf et al., 2013). It is likely that A-levels have a similar role in underpinning the academic backbone in psychology and other courses.

There is less evidence about degree performance in selection, although our UCL DClinPsy study found it predicted course exam performance independently from A-levels on Year 1, but not Year 2, exams (Scior, 2013).

Selecting on person-based factors
Interviews can take many forms and there is conflicting evidence as to their utility (Ferguson et al., 2002; Goho & Blackman, 2006; Hogg et al., 2014; Scior et al., 2013). Concerns have also been expressed about the reliability and fairness of traditional unstructured interviews (Prideaux et al., 2011; Salvatori, 2001).

The multiple mini-interview (MMI) is a practical structured interview increasingly used in medicine. Candidates rotate around ‘stations’ in order. At each station they perform a task – such as talking to an actor about an ethical issue – for a set time. An examiner at each station assesses performance using a structured checklist, to give an overall measure of non-cognitive ability (Dowell et al., 2012; Eva et al., 2004). The MMI is more reliable than unstructured interviews and measures actual rather than reported performance. In medicine, the MMI is a good predictor of performance in practical examinations of knowledge application and communication skills (Eva et al., 2009; Husbands & Dowell, 2013; Kelly et al., 2014). To our knowledge, there is no evidence as to the effectiveness of an MMI in clinical psychology.

Situational judgement tests (SJT) are written tests in which candidates are presented with hypothetical scenarios requiring them to choose or rank the most appropriate responses. SJTs have only recently been used in medicine – to select qualified doctors into specialty training (see Patterson et al., 2012, for a review). The SJT used to select candidates into general practice training has a small to medium correlation (r = .3) with job supervisor ratings a year later and a slightly higher correlation (r = .4) with the GP exit examination two years later,
and adds significantly to information provided from written knowledge and practical clinical tests (Patterson et al., 2013).

SJTs are not used to select medical undergraduates in the UK, although the UKCAT test is piloting one. In Belgium a study of 5000 medical school applicants found that a video-based SJT measuring interpersonal skills (candidates watched 30 videos and had decide what to do after each one) was a reasonable predictor of medical school assessments of interpersonal behaviour, and of performance as a doctor seven years later – correlations being small to medium (.15 to .23). It was significantly better than a ‘cognitive’ test of science knowledge and general mental ability, although the cognitive test was a better predictor of knowledge test scores (Lievens, 2013). As far as we know, there are no published reports of SJTs in clinical psychology selection, although anecdotally they are used or currently being piloted by several UK courses.

Aptitude tests in selection
The best-established aptitude test in medicine is the Medical College Admissions Test (MCAT), used in one form or another in the United States since the 1920s (Moss, 1930). MCAT has a small to medium correlation $(r = .4)$ with early medical school performance, and a medium to large correlation $(r = .6)$ with the US Medical Licensing Examination (similar to medical school finals in the UK) (Donnon et al., 2007). In predictive validity studies, a correlation of .6 is high. By contrast, there is only a small correlation $(r = .2)$ between UKCAT scores and medical school results, which is weaker than the correlation between school exam results and medical school results $(r = .3)$. UKCAT also adds little information over the school achievement mark (McManus, Dewberry et al., 2013a). BMAT has similar predictive power to UKCAT (McManus et al., 2011). Differences in the relative predictive power of aptitude tests reflect differences in their content. Until 2012 the MCAT subtests were Physical Science, Biological Science, Verbal Reasoning, and a Writing Sample. The best-performing subtest was Biological Science, which predicted medical school written knowledge test performance and job performance. The next best was Verbal Reasoning, which predicted job performance only (Donnon et al., 2007). UKCAT tests reasoning, not science knowledge. BMAT’s knowledge subscale performs significantly better than its reasoning and writing subscales (McManus et al., 2011).

Aptitude tests are not widely used in clinical psychology selection, and there is no consensus on what a psychological aptitude test should entail, although Baron (2011) suggests cognitive ability, research comprehension, personality, English language ability or SJTs might be useful. There are some course-specific tests, but there are insufficient data to assess their predictive validity reliably. From April 2015 all providers of NHS-funded training programmes in England are expected to assess candidates’ values and behaviours using reliable and valid methods, such as face-to-face structured interviews or multiple mini-interviews, and to use tools such as SJTs and aptitude tests to screen for interview selection (Health Education England, 2014). It is imperative to consider how these should be used to select in clinical psychology, where there is little research evidence.

Job performance
So far we have looked at whether selection methods predict course performance. While a student cannot qualify without passing, job performance is much more important. Selection methods are crucial because nearly everyone accepted on a medical or clinical psychology course qualifies. We could say this means selection methods work, but that assumes course assessments are proxies for job performance; that is, that candidates who perform well on the course go on to be good practitioners, candidates who perform badly go on to be poor practitioners, and candidates who fail the course would have been terrible practitioners.

Is this true? It is certainly controversial. Doctors often say that exams predict exams and they have nothing to do with being a good doctor (Hurwitz & Vass, 2002). To test this we need to look at how well the methods used to select on to courses predict job performance. To our knowledge, no direct evidence exists of the link between a psychological test and job performance, that chain of evidence would help validate selection methods.

One problem with establishing the chain is that there is no agreed definition of ‘good’ performance, and in the UK there are as yet few available data, on any potential measures of it. Being stuck off the medical register is, however, a reasonable proxy for bad performance, and there is evidence that medics who do badly in exams but scrape through are more likely to have licensing problems (Ludka et al., 2013; Papadakis et al., 2008) and more patient deaths (Norcini et al., 2014). To our knowledge, there are no studies of the link between course and job performance in clinical psychology.

Diversity and selection methods
Despite concerted efforts to widen access, the lack of diversity in the professions remains a problem (Cabinet Office, 2008). Newly qualified clinical psychologists are overwhelmingly female and white (Cape et al., 2008). In medicine, black and minority ethnic (BME) groups are better represented, but this is due to high proportions of some groups (e.g. Indian), while others (e.g. Bangladeshi and Black Caribbean) remain underrepresented (BMA, 2008). Fifty per cent of doctors are independently schooled compared with 6.5 per cent of British children (Cabinet Office, 2009).

35, 1056–1064.


Medical Schools Council (2014). Selection
School attainment is linked to gender, ethnicity, socio-economic status, and school characteristics (Crawford, 2014; Department for Education, 2014), which is clearly a problem and can mean able students do not apply for courses. For selectors, a major issue is whether it is possible to increase diversity without reducing quality, and one test of that is whether selection methods predict equally well for students from different groups.

Prior academic performance and diversity

It may disadvantage applicants from state schools to select on A-levels without any adjustment for school characteristics. Students from state schools perform slightly better at university than students with equivalent A-level grades from independent schools (Crawford, 2014). The effect disappears for those with the highest grades (HEFCE, 2014), from which the pool of medical students is drawn; however, the UKCAT-12 study did find that students from selective schools did slightly worse at medical school than those from state schools; and students who did particularly well in A-levels who would not usually be admitted to medical school. Students on the EMDP perform significantly worse in finals and have three times the failure rate of students on the conventional programme (McManus, Dewberry et al., 2013b). Students on the EMDP are different in many ways from those on the conventional programme, but these findings should encourage caution before lowering entry grades without research into how much they should be lowered, and for which groups.

On a related note, in the UCL DClinPsy study, graduates from Oxford or Cambridge performed best, and graduates from post-1992 institutions and non-UK universities performed worst on a statistics exam, with no differences found on all other assessments (Scior et al., 2013). Oxford and Cambridge medical graduates outperform academically those from other medical schools (McManus, Elder et al., 2008; Woolf et al., 2012). Bearing in mind that psychology undergraduates from BME backgrounds tend to be concentrated in post-1992 universities (Turpin & Fensom, 2004), further research is needed to understand the implications for selection into postgraduate courses.

Apptitude tests and diversity

Apptitude tests used in medicine, law and other subjects claim to more fairly assess ability than school exams (http://admissionstestingservice.org) yet better UKCAT and BMAT performance is predicted by grammar or independent school attendance, white ethnicity, having professional parents, being male and speaking English as a first language (Emery et al., 2011; Tiffin et al., 2014; McManus, Dewberry et al., 2013a). This suggests that, while not necessarily biased, aptitude tests are unlikely to increase diversity.

Interviews, situational judgment tests and diversity

The potential for bias in traditional interviews is well known. There is surprisingly little evidence about how the use of the MMI and SJTs impacts on diversity in selection in medicine or clinical psychology. Two recent studies from Australia and England found that MMI scores were unrelated to socio-economic factors, although applicants had already been screened for prior academic performance (Griffin & Hu, 2013; Taylor et al., 2013).

Course performance and diversity

There are concerns about the fairness of course assessments. In medicine, as in higher education generally, BME students and doctors, and those from lower socio-economic groups tend to perform less well (Woolf et al., 2011; Woolf et al., 2012). The UCL DClinPsy study found BME students were more likely to fail course assignments (Scior et al., 2013).
In medicine, numerous explanations put forward to explain the ethnic difference fail to do so (Woolf et al., 2012), but it is unlikely to result from direct discrimination by examiners or in the assessments themselves (McManus, 1996; Woolf et al., 2011). New avenues of investigation include negative stereotyping and stereotype threat, as well as how the relationship between ethnicity, friendship and belonging may influence the acquisition of ‘medical capital’ (Woolf et al., 2012).

Looking to the longer term
Being a psychologist, medic or another healthcare professional is a complex job requiring high-level knowledge and skills and appropriate attitudes and behaviours, with potentially life-changing effects for the population. The evidence for and against the variety of methods used to select suitable candidates is patchy, although some methods are clearly better than others.

There is good evidence that the wide use of A-levels to select onto undergraduate courses is justified by their ability to predict performance on academic and person-based outcomes. For postgraduate courses, limited evidence suggests that degree class is of some use in addition to A-levels. Aptitude tests that measure general factors such as reasoning add little or nothing to A-levels and GCSEs; tests measuring subject-specific knowledge are better. In selecting for personal qualities, traditional-style unstructured interviews should not be used. Highly structured practical tests such as the MMI are useful, although expensive and labour-intensive. SJTs seem promising, predicting both academic and job performance ratings in postgraduate medicine. More research is needed to establish what SJTs should look like in clinical psychology and for selection onto undergraduate courses.

When considering whether selection methods work, job performance post-qualification is crucial; however, there is little evidence. The assertion that high qualification rates are evidence of the quality of selection methods in choosing competent future practitioners is only justified if course performance predicts job performance. A handful of medical studies have found that poor course performance predicts being struck off the medical register and increase patient mortality rates. There is a considerable need for more longitudinal studies with job performance outcomes.

The professions of medicine and clinical psychology must do more to increase diversity, but the issues are complex. The relatively poor performance of some groups on A-levels, aptitude tests and SJTs may reflect real differences rather than bias per se, which could be better tackled by intervening much earlier in the educational pathway. A major issue for selectors is whether it is possible to adjust entry criteria to reduce selection bias, while maintaining standards. A-level entry criteria may be adjusted to ensure candidates from state and poorly performing schools are not disadvantaged, although this needs to be done by individual courses in a careful evidence-based way to avoid lowering standards. Adjusting for ethnicity is more difficult. There is no evidence that aptitude tests improve diversity. Interestingly, limited evidence suggests the MMI is perhaps fairer than other ways of assessing person-based factors.

In conclusion, we appear to be doing reasonably well at selecting candidates who will do well on courses with tests of knowledge, skills and attitudes. We have little idea, though, what this means in the longer term. Future research should focus on the relationship between selection and job performance, and a key issue is that there are not agreed, clear definitions as to what it is to be a ‘good’ practitioner.

The identification of suitable selection methods should be more coupled with research into what it means to be a safe and effective practitioner, rather than merely measures of whether one does well during training. Understanding how to remedy this situation very likely requires studies that go beyond specific courses and look at factors beyond exam performance both in terms of selection metrics and outcomes.

Note
The authors wish to note that the views expressed in this article are theirs alone and in no way represent the position of the UCL Clinical Psychology course or UCL Medical School.

"Future research should focus on the relationship between selection and job performance"
Keeping a spotless mind: The neuroscience of ‘motivated forgetting’

FREE EVENT
Thu 17 Sep
6–7.15pm

Dr Michael Anderson, University of Cambridge
Join Dr Anderson as he discusses research revealing how the brain accomplishes motivated forgetting, and how these mechanisms shape what we remember of life, protecting our mental health.

FIND OUT MORE: www.britishacademy.ac.uk/events

In partnership with

BRITISH ACADEMY
for the humanities and social sciences

The British Psychological Society

10–11 Carlton House Terrace, London SW1Y 5AH

Charing Cross / Piccadilly