

# Fair and equal assessment in postgraduate training?

A future research agenda

Assessment of GP training has historically shown differences in performance according to ethnicity and sex.<sup>1</sup> A recent meta-analysis<sup>2</sup> shows that ethnic differences in attainment are a consistent feature of medical education in the UK, 'being present across medical schools, exam types, and undergraduate and postgraduate assessments, and have persisted for at least the past three decades'. The most substantial differences are found for doctors taking postgraduate examinations as international medical graduates, who frequently under-perform compared to their locally-trained peers.<sup>3</sup>

Although these differences have been repeatedly found, we lack a satisfactory explanation that can guide what may be done to address the issues. In this editorial, we propose an agenda to guide future research and to encourage debate within the broader scientific and academic community. Using an interdisciplinary approach, we draw on expert inputs from a variety of academic and stakeholder experiences to summarise key issues surrounding fairness in assessment. These opinions were gathered at a 1-day seminar at the University of Cambridge.

There exists understandable concern amongst key stakeholders regarding the threat of litigation when research findings expose unexplained significant group differences in assessment outcomes. We argue that postgraduate assessment methodologies are not inherently 'biased'. The reasons for observed group differences are far more complex, and in need of far more systematic research, leading to more open dialogue between all stakeholders.

### FUTURE RESEARCH DESIGN

The use of innovative research designs to explore causality in observed group differences is recommended. New approaches should include:

- adopting interdisciplinary research perspectives to encourage fresh thinking from several vantage points;
- longitudinal tracking of trainees from medical school through to end-of-training to include more granular information (for example, an individual's educational history); and
- developing appropriate theoretical frameworks to interpret findings, using contemporary research in adult intellectual development and learning.

New research methodologies could provide original insights such as: (1) social network analysis in educational settings; (2) structural modelling to explore causality over the training pathway; (3) further large scale meta-analytic studies; (4) sociolinguistic analyses of discourse in intercultural assessment encounters; (5) retrospective accounts from diary studies; and (6) combining quantitative with qualitative research designs to explore the meaning behind patterns of data emerging. Using new research designs and methodologies we propose focusing on four key research areas.

### 1. CLASSIFICATION SYSTEMS FOR ETHNIC GROUPS

A major problem with existing research is that the use of classification systems regarding ethnic groups is often muddled and does not take into account to how ethnic categories are differently understood within different theoretical frameworks. For example, the term 'Asian' was not in any general use until Idi Amin's expulsion of all 'Asians' from Uganda in 1972. Few doctors classified as 'Asian' in the UK would recognise and identify themselves socially as 'Asian'. Even narrower specifications, for example, South Asian, are inadequate as such geographical areas cover a wide

range of ethnic groupings and pay no attention to how ethnicity plays a role in identity and performance. An improved terminology to classify data precisely in order to allow more accurate interpretation of findings is required.

### 2. ASSESSMENT CRITERIA AND TEST MODALITIES

There is much to be gained from investigating how linguistic practices involved in assessment are communicated and understood by all those involved. It could be argued that the language used to describe marking schemes is not culturally neutral. We argue for a broader range of test modalities to be considered, as all assessments are also culturally located.

Do we have a unitary concept of what makes an excellent doctor? If so, is that description gendered or ethnicised in some manner? These conceptualisations must also be reflected in the curriculum, thus ensuring there are opportunities for people to excel. A 'hidden curriculum' may exist where who gets help and how someone asks for it affects their learning opportunities. Assumptions about the most appropriate behaviour are transmitted in the teaching. A sociocultural analysis of approaches to teaching is a worthy avenue for further research, exploring how learning contributes to practice over time.

### 3. FACTORS THAT INFLUENCE CAREER CHOICES

Researchers refer to the 'fallacy of the single factor', as they acknowledge the multifaceted nature of issues explaining group differences in assessment. In future, multifactorial analyses are required, to include variables ranging from the psychological (personality factors, such as conscientiousness), and the socioeconomic (for example, parents' earnings, access to education), to the cultural (for example, one's position in a social network and the social capital that provides).

Detailed case studies focusing on outliers and/or extreme groups could provide a particularly rich source of data in understanding how these factors operate. For example, there are women who consistently outperform males on educational assessments, and there are those from ethnic minority groups who

---

*"Do we have a unitary concept of what makes an excellent doctor? If so, is that description gendered or ethnicised in some manner? These conceptualisations must also be reflected in the curriculum, thus ensuring there are opportunities for people to excel."*

---

---

*“Seeking to counter this systematic unintended discrimination could be the single most important way of ensuring the highest standards of training.”*

---

perform significantly better than normative considerations would suggest. Understanding the detail of such cases may also suggest ways in which apparent barriers to success may be overcome.

#### 4. IMPLICATIONS OF OBSERVED GROUP DIFFERENCES FOR PRACTICE

A richer understanding of the implications of the observed groups differences in practice is needed. Research here is essential for understanding the issues for the individuals involved, and for the social systems within which these observed findings occur. Results from a recent PMETB survey (PMETB was the Postgraduate Medical Education and Training Board in the UK, now subsumed in the General Medical Council) show that those from ethnic minority groups reported errors in their practice significantly less frequently compared to their UK/Irish white counterparts.<sup>4</sup> What are the implications for patient safety? Is fear of retribution greater for some groups or could the social norms of how one approaches work be defined differently between ethnic groups, without consequences for outcomes in practice?

#### CONCLUDING COMMENTS

Four key areas to guide further research are presented, ranging from design issues to analysing outcomes in practice. This list is not exhaustive, and many issues surrounding making assessment fair and equal remain under-researched. Evidence suggests that assessments in

postgraduate training do not necessarily systematically exhibit bias by underpredicting the performance of minority group members. (Note this is not an endorsement of all assessment tools; there are likely to be poorly designed assessments in existence.) This is an important conclusion as it directs researchers to explore factors other than biased assessment tools as the major determinant of group differences.

In practice, the combination of selection and training placement systems often operate against the interests of the weaker recruits (that is, those candidates performing least well at selection are assigned to the least popular training placements, thereby encouraging a cycle of educational deprivation). Seeking to counter this systematic unintended discrimination could be the single most important way of ensuring the highest standards of training.

Educational institutions and health service organisations value a diverse workforce, so identifying causes of group differences and developing effective interventions is a serious challenge. Further research is required to deliver best practice guidelines for future design of assessments in postgraduate medical education and training. It will help Royal Colleges and other testing bodies to develop the sophistication of their responses to what, on the face of it, may present as distasteful outcomes of their assessments.

#### ADDRESS FOR CORRESPONDENCE

##### Fiona Patterson

City University, Northampton Square, London, EC1V 0HB, UK.

E-mail: F.patterson@workpsychologygroup.com

##### Fiona Patterson

Professor, University of Cambridge, Cambridge and Work Psychology Group, Ashbourne.

##### Mei-Ling Denney

Interim MRCGP Research & Development Lead for the MRCGP, Edinburgh.

##### Richard Wakeford

University of Cambridge, Cambridge.

##### David Good

University of Cambridge, Cambridge.

##### Provenance

Commissioned; not externally peer reviewed.

DOI: 10.3399/bjgp11X612936

#### REFERENCES

1. Wood R. *Assessment and testing: a survey of research*. Cambridge: Cambridge University Press, 1991: 166–183.
2. Woolf K, Potts HW, McManus IC. Ethnicity and academic performance in UK trained doctors and medical students: systematic review and meta-analysis. *BMJ* 2011; **342**: d901.
3. Eva KW, Wood TJ, Riddle J, et al. How clinical features are presented matters to weaker diagnosticians. *Med Educ* 2010; **44**(8): 775–785.
4. PMETB. *National Training Surveys 2008–2009 Key findings*. London: Postgraduate Medical Education Training Board, 2010: 51.