Work Psychology Group

Thinking differently

Selecting the “right” candidates for the health professions: why and how?

Prof Fiona Patterson

Work Psychology Group

Overview

- **Context of selection** in the healthcare professions
- Research evidence on the **quality of different selection methods**
  - Case Study: **Medical & dental school admissions using UKCAT**
- Implications for **policy & practice**
- **Current** & **future research**

Yale Medical School graduating class of 1924

Yale University, Harvey Cushing/John Hay Whitney Medical Library

An emphasis on values in recruitment

- **Compassion, benevolence, respect & dignity**, are important for any healthcare professional to ensure high quality care
- Identified examples of failure in care, lacking **compassion & competence**, a broken **organisational culture**, **workload pressures** & staff without the **right values**
- Amongst many other initiatives Health Education England are concerned to assess **values in recruitment**

Context of Selection in Healthcare Today

- Internationally, selection continues to be **highly competitive**
- Practically, this means selection can be **resource intensive**
- **High-stakes** selection context
- Selection frequently driven by **political considerations & stakeholder acceptance**
- **Risk of litigation** if the method is viewed as 'unfair'
- Increased emphasis on **widening access** & the importance of **non-academic attributes**
- Pressing need to address **workforce shortages** in some specialties/remote & rural locations

Some key questions

What attributes are important to be an effective clinician, now and in the next 25 years?

What selection methods are available to test these accurately?

Given the costs - beyond some basic assessment – is a lottery the best option?

Why not use a lottery system?

Dutch medical schools abandon selection for lottery system for places

Jan Coebergh *Newcastle*

Two Dutch medical schools will no longer independently select some students as these students do not perform better at medical school.

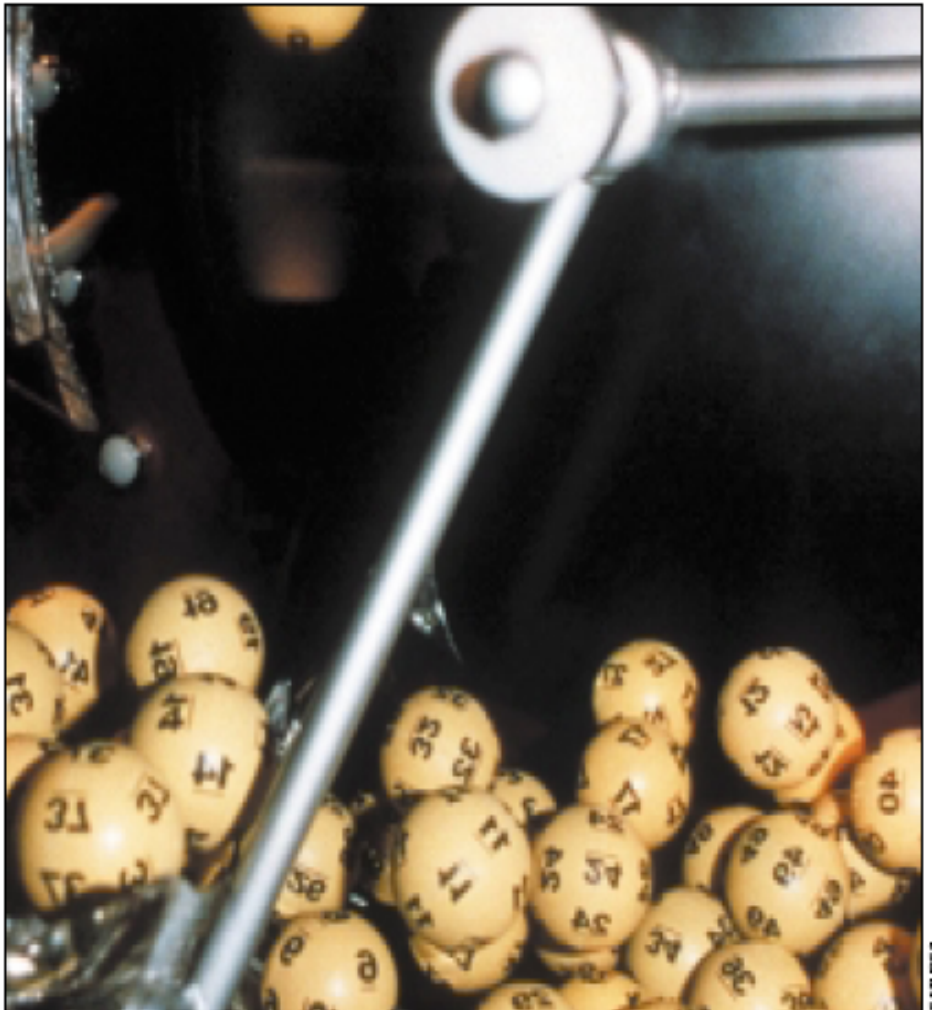
Until 1999, admission into the nine Dutch medical schools was based entirely on a lottery. Based on academic grades, the average chance of getting a place was 35%, rising to 70% for those with the highest grades.

The law changed in 1999 after a media row over a bright girl who was not allowed to enter medical school three years in a row. Universities were allowed to select up to half their

Some used this allowance to promote entry of mature students, graduates, and ethnic minorities.

In recent evaluations at four universities, three found that selected students did not get higher grades than those given places by the lottery. They concluded that selection was not beneficial. Two universities will stop selecting since the costs are high and will return to the lottery admission policy. One university did find that selected students performed better and will continue.

In contrast to the experiences



How effective are selection methods in medical education?

Patterson, Knight, Dowell, Nicholson, Cousans, Cleland (2016) Medical Education.

- Academic records
- Aptitude tests
- Personal statements
- References
- Situational Judgement Tests
- Personality assessement
- Interviews & MMIs

How effective are selection methods in medical education? A systematic review

Fiona Patterson,[1] Alec Knight,[2] Jon Dowell,[3] Sandra Nicholson,[4] Fran Cousans[2] & Jennifer Cleland[5]

CONTEXT Selection methods used by medical schools should reliably identify whether candidates are likely to be successful in medical training and ultimately become competent clinicians. However, there is little consensus regarding methods that reliably evaluate non-academic attributes, and longitudinal studies examining predictors of success after qualification are insufficient. This systematic review synthesises the extant research evidence on the relative strengths of various selection methods. We offer a research agenda and identify key considerations to inform policy and practice in the next 50 years.

METHODS A formalised literature search was conducted for studies published between 1997

(vii) interviews and multiple mini-interviews (MMIs), and (viii) selection centres (SCs). The evidence relating to each method was reviewed against four evaluation criteria: effectiveness (reliability and validity); procedural issues; acceptability, and cost-effectiveness.

CONCLUSIONS Evidence shows clearly that academic records, MMIs, aptitude tests, SJTs and SCs are more effective selection methods and are generally fairer than traditional interviews, references and personal statements. However, achievement in different selection methods may differentially predict performance at the various stages of medical education and clinical practice. Research into selection has been over-reliant on cross-

Selection Method	Reliability	Validity	Candidate acceptability	Promotes widening access?
Academic records				
Structured Interviews/MMIs				
Situational Judgement Tests				
Aptitude testing				
Personality Tests				
Traditional Interviews				
Personal statements				
References				

Patterson, et al, 2016. How effective are selection methods in medical education and training? A systematic review. Medical Education.

Selection Method	Reliability	Validity	Candidate acceptability	Promotes widening access?
Academic records	High	High	High	Low
Structured Interviews/MMIs	Moderate to high	Moderate to high	High	Moderate
Situational Judgement Tests	High	High	Moderate to high	High
Aptitude testing	High	Various	Moderate	Moderate
Personality Tests	High	Moderate	Low to moderate	N/A
Traditional Interviews	Low	Low	High	Low
Personal statements	Low	Low	High	Low
References	Low	Low	High	Low

Patterson, et al, 2016. How effective are selection methods in medical education and training? A systematic review. Medical Education.

Selection Method	Reliability	Validity	Candidate acceptability	Promotes widening access?
Academic records	High	High	High	Low
Structured Interviews/MMIs	Moderate to high	Moderate to high	High	Moderate
Situational Judgement Tests	High	High	Moderate to high	High
Aptitude testing	High	Various	Moderate	Moderate
Personality Tests	High	Moderate	Low to moderate	N/A
Traditional Interviews	Low	Low	High	Low
Personal statements	Low	Low	High	Low
References	Low	Low	High	Low

Patterson, et al, 2016. How effective are selection methods in medical education and training? A systematic review. Medical Education.

Selection Method	Reliability	Validity	Candidate acceptability	Promotes widening access?
Academic records	High	High	High	Low
Structured Interviews/MMIs	Moderate to high	Moderate to high	High	Moderate
Situational Judgement Tests	High	High	Moderate to high	High
Aptitude testing	High	Various	Moderate	Moderate
Personality Tests	High	Moderate	Low to moderate	N/A
Traditional Interviews	Low	Low	High	Low
Personal statements	Low	Low	High	Low
References	Low	Low	High	Low

Patterson, et al, 2016. How effective are selection methods in medical education and training? A systematic review. Medical Education.

Selection Method	Reliability	Validity	Candidate acceptability	Promotes widening access?
Academic records	High	High	High	Low
Structured Interviews/MMIs	Moderate to high	Moderate to high	High	Moderate
Situational Judgement Tests	High	High	Moderate to high	High
Aptitude testing	High	Various	Moderate	Moderate
Personality Tests	High	Moderate	Low to moderate	N/A
Traditional Interviews	Low	Low	High	Low
Personal statements	Low	Low	High	Low
References	Low	Low	High	Low

Patterson, et al, 2016. How effective are selection methods in medical education and training? A systematic review. Medical Education.

What are Situational Judgement Tests?

- Situational Judgement Tests (SJTs) are a **measurement method** designed to assess judgement in role-relevant situations:
  - Present challenging situations likely to be encountered in the role
  - Candidates make judgements about possible responses
  - Scored against pre-determined key
- SJTs focus on **non-academic attributes** *(e.g. integrity, empathy, resilience, team involvement)*

Example SJT item (for entry into postgraduate training)

You are reviewing a routine drug chart for a patient with rheumatoid arthritis during an overnight shift. You notice that your consultant has inappropriately prescribed methotrexate 7.5mg daily instead of weekly.

Rank in order the following actions in response to this situation (1=Most appropriate; 5=Least appropriate)

A Ask the nurses if the consultant has made any other drug errors recently
B Correct the prescription to 7.5mg weekly
C Leave the prescription unchanged until the consultant ward round the following morning
D Phone the consultant at home to ask about changing the prescription
E Inform the patient of the error

2015, 1–15, Early Online MEDICAL TEACHER

AMEE GUIDE

Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100

FIONA PATTERSON[1,2], LARA ZIBARRAS[3] & VICKI ASHWORTH[1]

[1]Work Psychology Group, UK, [2]University of Cambridge, UK, [3]City University London, UK

Abstract

Why use SJTs? Traditionally, selection into medical education professions has focused primarily upon academic ability alone. This approach has been questioned more recently, as although academic attainment predicts performance early in training, research shows it has less predictive power for demonstrating competence in postgraduate clinical practice. Such evidence, coupled with an increasing focus on individuals working in healthcare roles displaying the core values of compassionate care, benevolence and respect, illustrates that individuals should be selected on attributes other than academic ability alone. Moreover, there are mounting calls to widen access to medicine, to ensure that selection methods do not unfairly disadvantage individuals from specific groups (e.g. regarding ethnicity or socio-economic status), so that the future workforce adequately represents society as a whole. These drivers necessitate a method of assessment that allows individuals to be selected on important non-academic attributes that are desirable in healthcare professionals, in a fair, reliable and valid way.

What are SJTs? Situational judgement tests (SJTs) are tests used to assess individuals' reactions to a number of hypothetical role-relevant scenarios, which reflect situations candidates are likely to encounter in the target role. These scenarios are based on a detailed analysis of the role and should be developed in collaboration with subject matter experts, in order to accurately assess the key attributes that are associated with competent performance. From a theoretical perspective, SJTs are believed to measure prosocial Implicit Trait Policies (ITPs), which are shaped by socialisation processes that teach the utility of expressing certain traits in different settings such as agreeable expressions (e.g. helping others in need), or disagreeable actions (e.g. advancing ones own interest at others, expense).

Are SJTs reliable, valid and fair? Several studies, including good quality meta-analytic and longitudinal research, consistently show that SJTs used in many different occupational groups are reliable and valid. Although there is over 40 years of research evidence available on SJTs, it is only within the past 10 years that SJTs have been used for recruitment into medicine. Specifically, evidence consistently shows that SJTs used in medical selection have good reliability, and predict performance across a range of medical professions, including performance in general practice, in early years (foundation training as a junior doctor) and for medical school admissions. In addition, SJTs have been found to have significant added value (incremental validity) over and above other selection methods such as knowledge tests, measures of cognitive ability, personality tests and application forms.

Case Study.

Using SJTs for medical & dental school admissions

N= 26,000 per year for 8,000 posts

5 subtests

- Verbal, numerical, abstract reasoning & decision analysis
- **SJT – targets empathy, integrity & team involvement**

http://www.ukcat.ac.uk/

Test Specification

- An SJT for a novice population (**no clinical knowledge required**)

Content

- Scenarios based in either a healthcare setting or during education/training for a medical/dental career
- Third party perspective

Response Format (rating using a 4 point scale)

- Rate the ***appropriateness*** of a response from ‘very appropriate’ to ‘very inappropriate.
- Rate the ***importance*** of a response from ‘very important’ to ‘not important at all’

Example UKCAT SJT items *(appropriateness)*

A consultation is taking place between a senior doctor and a patient; a medical student is observing. The senior doctor tells the patient that he requires some blood tests to rule out a terminal disease. The senior doctor is called away urgently, leaving the medical student alone with the patient. The patient tells the student that he is worried he is going to die and asks the student what the blood tests will show.

*How **appropriate** are each of the following responses by the medical student in this situation?*

Q1 Explain to the patient that he is unable to comment on what the tests will show as he is a medical student

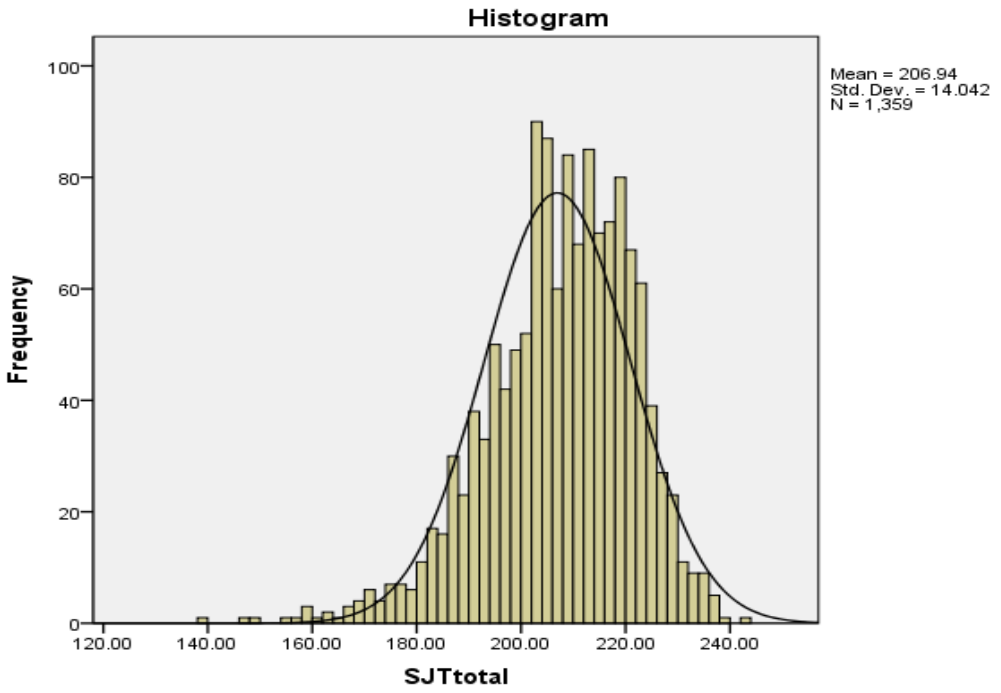
Q2 Acknowledge the patient's concerns and ask whether he would like them to be raised with the senior doctor

Q3 Suggest to the patient that he poses these questions to the senior doctor when he returns

Q4 Tell the patient that he should not worry and that it is unlikely that he will die

UKCAT SJT Evaluation

- **Reliability** of a 70 item test with similar quality items estimated (α=.75 to .85)
- Candidate reactions shows **good face validity** (significantly more than the cognitive tests of UKCAT)
  - *Content of SJT relevant for med/dental applicants = 70%*
  - *Content of the SJT is fair to med/dental applicants = 63%*



Widening access in selection using situational judgement tests: evidence from the UKCAT

Filip Lievens,[1] Fiona Patterson,[2] Jan Corstjens,[1] Stuart Martin[3] & Sandra Nicholson[4]

CONTEXT Widening access promotes student diversity and the appropriate representation of all demographic groups. This study aims to examine diversity-related benefits of the use of situational judgement tests (SJTs) in the UK

and other minority ethnic candidates), and gender.

RESULTS Firstly, the effect size for SES was lower for the SJT (d = 0.13–0.20 in favour of the higher SES group) than it was for the cognitive tests (d = 0.38–0.35). Secondly, effect sizes for ethnicity of the SJT and cognitive tests were similar (d = –0.50 in favour of White candidates). Thirdly, males outperformed females on cognitive tests, whereas the reverse was true for SJTs. When equal weight was given to the SJT and the cognitive tests in

at less of a disadvantage and may potentially diversify the student intake. However, use of the SJT applied in this study did not diminish the role of ethnicity. Future research should examine these findings with other SJTs and other tests internationally and scrutinise the causes underlying the role of ethnicity.

"SJTs ….complement cognitive (academic) tests….puts candidates of lower socioeconomic status at less of a disadvantage & can diversify the student intake…"

Medical Education, 2016

Medical Education 2016; 50: 624–636
doi: 10.1111/medu.13060

Summary

- Good evidence emerging about the quality of different selection methods
- Differences in predictive validity of selection methods in **education vs in clinical practice**
- **Criterion problem** -*"what is a competent healthcare practitioner?"*
- Non-academic attributes should be used for '**selecting out**' & academic attributes used for '**selecting in**'
- Increased focus on selection tools/systems that promote **widening access** & **diversity** & that can **address workforce shortages**
- Some attributes are not currently assessed e.g. **creativity & innovation** *(Patterson & Zibarras, 2017)*
- How do we best use information at selection to **inform early education & training interventions?**

Forthcoming 2018; Patterson & Zibarras (Eds)

Chapter contributions from various authors:

Selection system design, evaluation approaches; Role analysis & defining selection criteria; Attraction & recruitment; Aptitude Tests; SJTs; Personality assessment; Interviews; Values based recruitment; Pro-sociality & altruism; Widening access & diversity issues; Candidate & stakeholder perceptions; Coaching issues; Future issues in research and practice

Combination of latest theory developments, research evidence alongside international case studies

Thank you

*www.**researchgate**.net/profile/**Fiona_Patterson/publications*

f.patterson@workpsychologygroup.com

What are SJTs measuring?

- SJTs measure **implicit trait policies** (**ITPs**) which are **beliefs about costs/benefits** of expressing certain traits (which guides behaviour)

- ITPs shaped by experiences in **socialisation processes** (e.g. early on in parental modelling) that teach the utility of;
  - agreeable expressions *e.g. helping others in need, turning the other cheek, looking after one's neighbours or*,
  - disagreeable actions *e.g. showing selfish preoccupation with one's own interests, holding a grudge/getting even, and advancing one's own interests at others' expense*

Group differences & content validity

- **SJT correlates with other subtests** (approx *r*=0.28). Since a large amount of variance is not explained, the SJT is assessing different constructs to the other tests.
- **Gender:** Females outperformed males (0.2 SD)
- **Ethnicity:** White candidates performed better (0.3SD)
- **Occupation & Employment Status:** those in the higher occupational classes (i.e. Managerial/Professional Occupations) do not always score higher than those in lower classes - in some cases those from lowest occupational groups, received the highest mean score.

Academic Records

- Most widely used selection method
- Potential bias against ‘non-traditional’ candidates

Strengths	Limitations
Good predictor of performance in education	Less predictive of clinical practice
Research is generally highly consistent	In the UK, A Levels are losing discriminating power
Generally administered by other bodies, so low cost to educators	Potential socio-economic class bias
Standardised and well-recognised assessments	

Aptitude Tests

- Mixed findings, depending on the specific aptitude test used (e.g. MCAT/ GAMSAT/ UKCAT/ BCAT/ UMAT/ HPAT)
- The broad range of tests available makes commenting on generality of findings problematic
- It is important to evaluate each aptitude test in their own right in order to draw conclusions regarding the quality of the tool

Strengths	Limitations
Some evidence for reliability and validity (incremental, predictive, criterion-related)	Reliability and validity may be affected by how they are used (i.e. weighting, cut score, etc)
	No evidence on cost-effectiveness at present
	May be less equitable for non-traditional applicants (e.g. SES)

Personal Statements

- Lack of good quality research in relation to this selection method

Strengths	Limitations
Popular & widespread selection method	Less reliable than other methods
Some evidence of predictive validity in relation to student performance	Potential data contamination caused by external influences (e.g. length of time to complete application, third party influence, location, *etc*.)
Some evidence of candidate acceptance	Potential for plagiarism
	Information inconsistently used during the decision-making process
	Highly resource-intensive to mark

References

- Little research examining the use of referee reports

Strengths	Limitations
Prevalent selection method in healthcare education	Clear consensus in the literature that poor predictors of performance
	Information contained in reports may result in admission bias
	Information inconsistently used by during the decision-making process
	Varied candidate perceptions of acceptability
	Highly resource-intensive to mark

Personality & Emotional Intelligence (EI) Assessment

- Overall variable quality of research for this selection method

Strengths	Limitations
Common assessment method used across multiple industries	No significant link between some personality tools & performance (e.g. MBTI, *etc*.)
Moderate to significant links between some personality traits and performance (e.g. PQA, NEO-PI-R, *etc*.)	Certain traits may be associated with different levels of performance over the course of healthcare education (e.g. conscientiousness and pre-/clinical performance)
Traits linked to performance: ▪ Empathy and motivation → positive ▪ 'Dysfunctional' personality → negative	Could lead to a reduction in the diversity of personalities amongst students
	Insufficient evidence to support use of EI in selection at present

Situational Judgement Tests (SJTs)

- High quality research, including meta-analyses/systematic reviews

Strengths	Limitations
An increasingly popular method of assessment in healthcare	Method of construction & response instructions may affect validity
Strong predictor of job performance; also predicts performance *above* cognitive ability & personality tests	Mode of administration may affect candidate reactions (e.g. computer-based vs. video-based)
Positive candidate reactions	Some item types may be more susceptible to faking, practice & coaching effects than others
Evidence that coaching does not significantly impact on validity	Requires expertise to design effectively
Reliable method of assessment with low adverse impact to minorities	

Interviews & Multiple Mini Interviews (MMIs)

- Widely used for many years
- Format varies widely – 'traditional', structured and MMI
- MMI increasingly popular, but design & implementation varies hugely

Strengths	Limitations
Means of assessing non-academic skills	Careful design is required to ensure good reliability
Good approach for some aspects, such as communication skills	Potential for bias (gender, ethnicity, SES)
High face validity	Resource intensive
Some evidence they can be ranked effectively	Rarely clear what content is actually assessed within a composite total score, especially with MMIs
Belief may help screen out 'unsuitable' entrants	Historically little evidence of predictive validity, though changing as interviews become more structured

A model for future design & evaluation of selection