Work Psychology Group
Thinking differently

Breaking the class ceiling:
Can situational judgement testing support widening access to the professions?

Professor Fiona Patterson

Overview

- The case for ‘breaking the class ceiling’?
- *Case Study 1:* ***Widening access in selection in medical & dental schools admissions using UKCAT***
- *Case Study 2:* ***Promoting diversity in selection in the banking sector***
- Implications for future research, theory & practice

The case for diversity & inclusion?

- Diversity confers a **competitive advantage**
  - Increased potential for **innovation** & **improved decision making**
  - Search for **top talent** draws from the widest possible pool
  - Customer service: **reflecting the communities served**
- Diversity as an '**organisational health**' indicator
- Fairness, **social justice** & **corporate social responsibility**
- Often embedded within the **organisation's values**

Laura Spence

- Laura Spence applied for medicine at Oxford having taken 10 GCSEs, obtaining the top A* grade in each.
- Spence was not offered a place because *"other candidates had equally good qualifications had performed better at interview"*
- Huge political row that Oxford had discriminated against her because of her state-school background in a "working-class" region
- Spence won a scholarship at Harvard to study biochemistry & later graduated in medicine from Cambridge
- The rejection of a well-qualified state-school pupil led to suspicions that Spence's exclusion was on the basis of **social class & regional prejudice** rather than academic suitability

How can we best design selection methods & systems to promote diversity in SES?

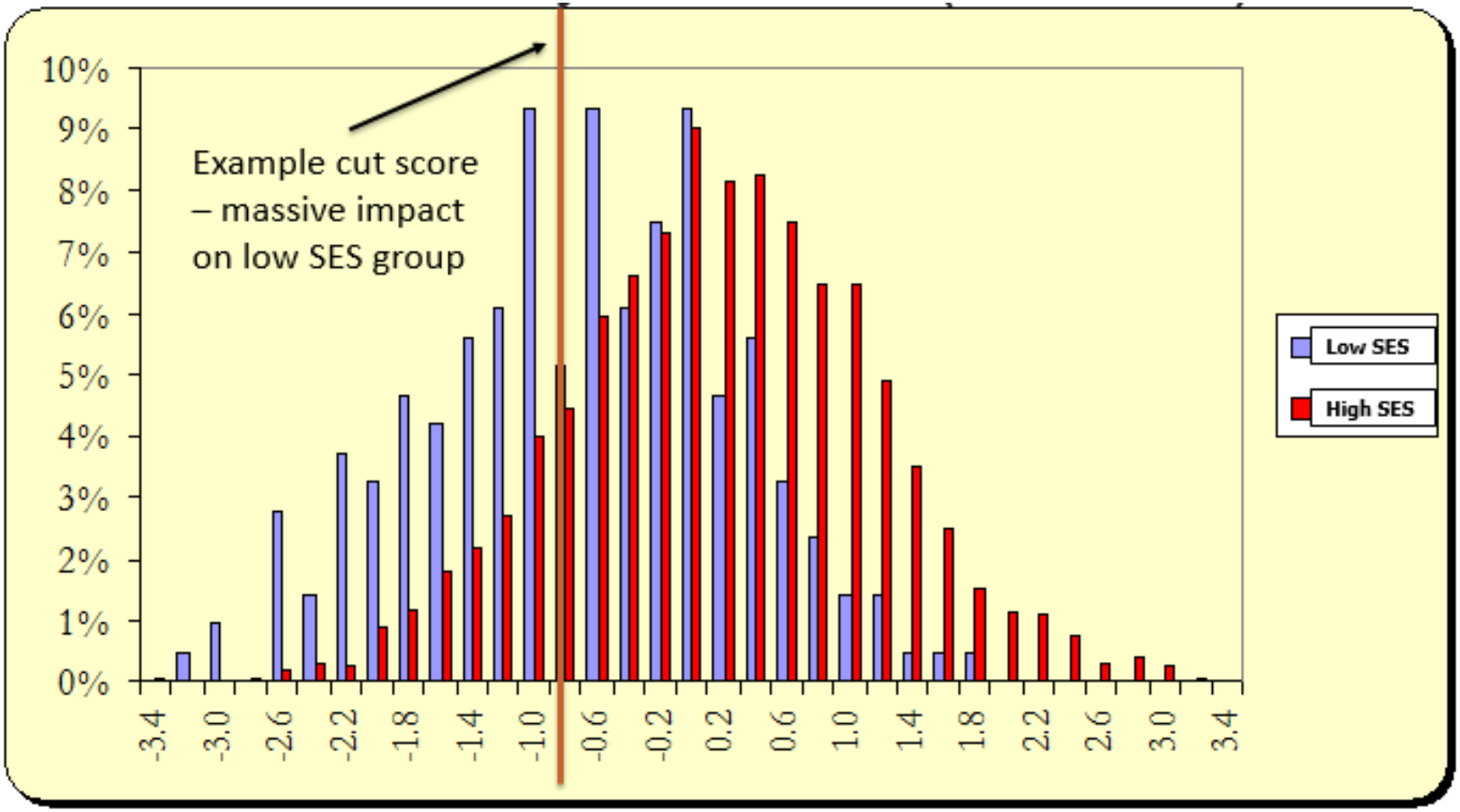
- Research tends to focus on outreach, attraction, candidate preparation, rather than **selection methods** (Ashley et al, 2016)
- Challenges & issues in assessment
  - **Defining SES?**
  - **Differential academic attainment** problem - lower SES is linked to lower academic achievement & slower rates of academic progress compared with higher SES communities (APA, 2016)
  - **Cognitive ability testing?**
  - Assessor (unconscious) **bias?**
  - Use of **contextual data?**

High volume selection methods: A levels

- 'Traditional' high-volume selection methods, e.g. **cognitive tests/A-levels, are increasingly incongruent** with a social mobility agenda
- Independent school pupils more than twice as likely as pupils in state schools to be accepted into one of the 30 most highly selective universities (Sutton Trust, 2016), introducing immediate bias in selection (Kirkup et al., 2008)
- **30%** of pupils from private schools gain 3 A's, compared to **10.7%** of pupils attending state schools (Paton, 2012)
- Private school students do not outperform state school students for **undergraduate degree class** (Smith & Naylor, 2001)
- Links between A level attainment & career success remain unclear (Kirkup et al., 2008)

Cognitive Ability Tests & SES

- Clear links between cognitive ability & job performance but negative impact on SES



Selection methods for the healthcare professions

Research evidence & practice

Work Psychology Group
Thinking differently

Medical Education, Jan 2016

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.

Evaluating the potential for UKCAT to promote diversity

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

5 subtests

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

UK Clinical Aptitude Test

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

Longitudinal assessment of the impact of the use of the UK clinical aptitude test for medical student selection

Jonathan Mathers, Alice Sitch & Jayne Parry

CONTEXT Medical schools are increasingly using novel tools to select applicants. The UK Clinical Aptitude Test (UKCAT) is one such tool and measures mental abilities, attitudes and professional behaviour conducive to being a doctor usin
affected by so
ditional meas
free to use U
broad modali
line', 'factor'
to provide th
assessing the
UKCAT on m
different soci

METHODS M
model the ou
cal schools during the period 2004–2011 (data obtained from UCAS), adjusted for sex, ethnicity, schooling, parental occupation, educational attainment, year of application and UKCAT use (borderline, factor and threshold).

RESULTS The three ways of using the UKCAT did not differ in their impact on making the selection process more equitable, other than a marked reversal for female advantage when applied in a 'threshold'
ongitu-
T in its
ersal of
strate
ns of
tive
rate
female
sed on
d
selective schooling. In view of this, the utility of the UKCAT as a means to widen access to medical schools among non-White and less advantaged applicants remains unproven.

"Our findings demonstrate no changes in admission rates based on higher social class…the (cognitive ability tests) are not a means to widen access to medical schools among less advantaged applicants"

SJT Specification

- An SJT for a novice population (**no medical 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

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

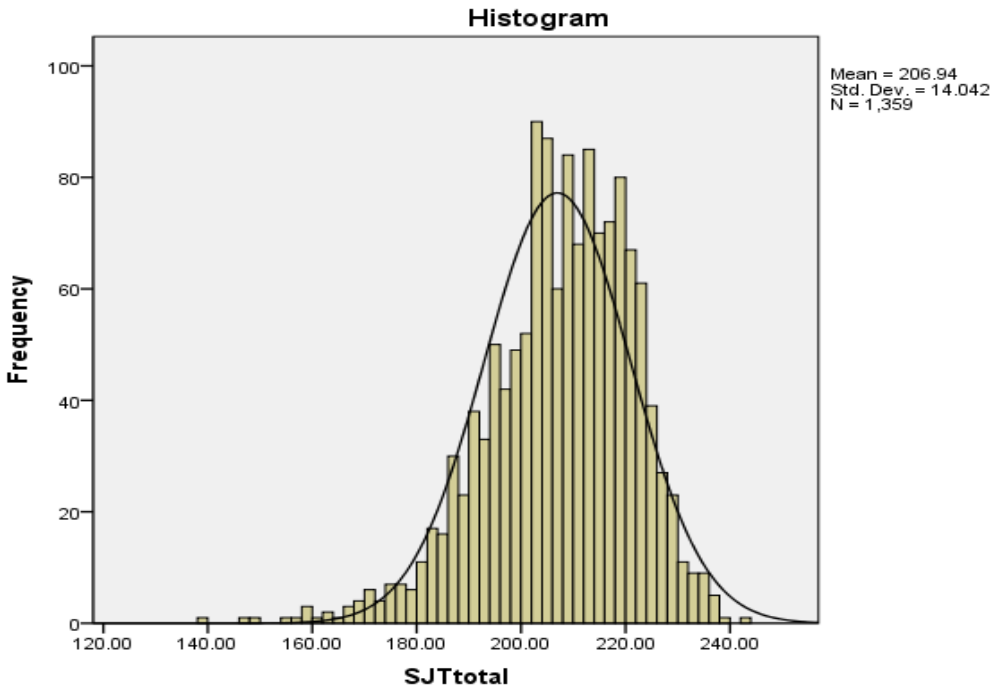
Q2 Acknowledge the patient's concerns and ask whether he would like them to be raised with the senior doctor

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

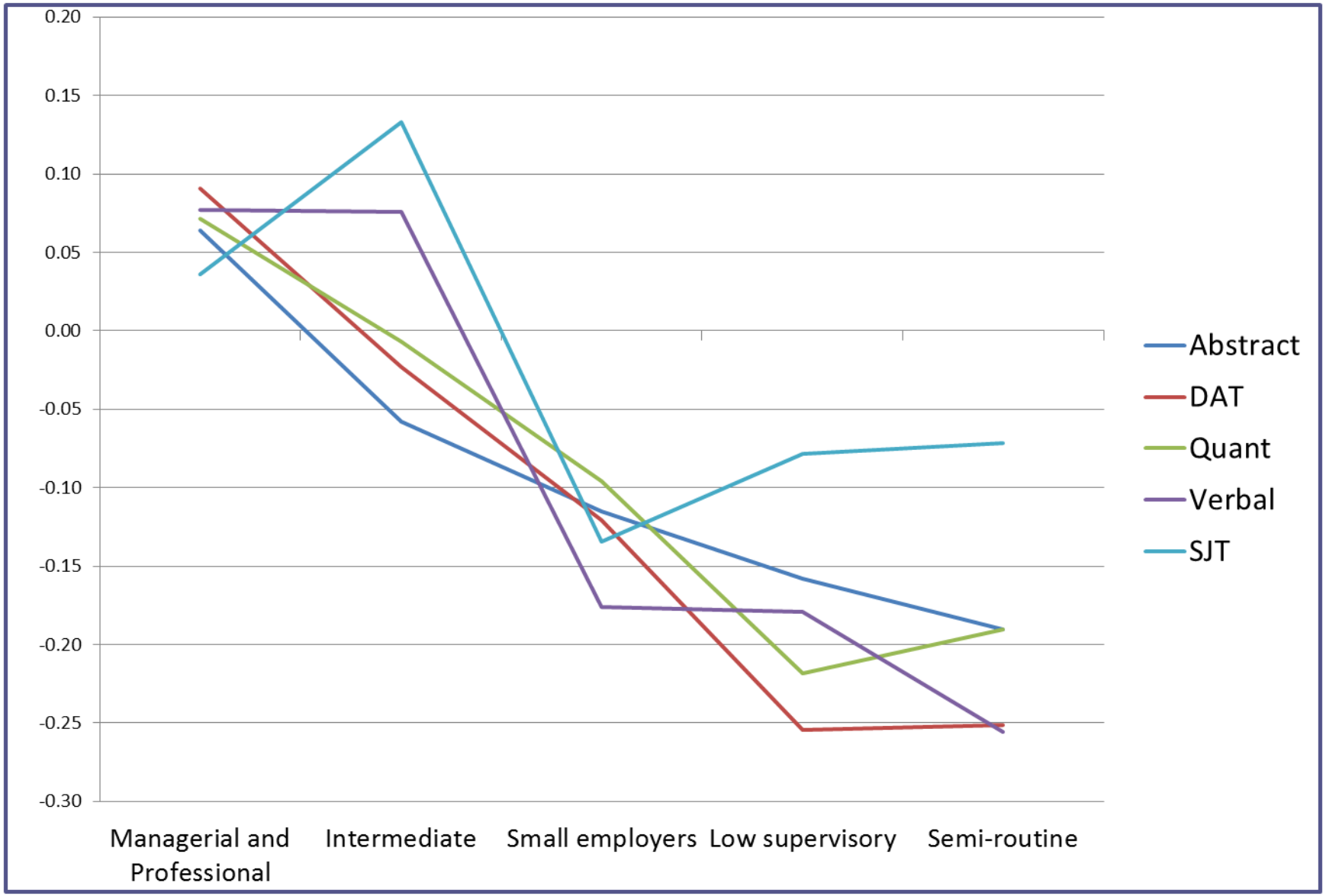
UKCAT SJT Evaluation

- **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%*



UKCAT SJT Evaluation

- **SJT correlates with CAT** (approx *r*=0.28). Since a large amount of variance is not explained, the SJT is assessing different constructs to the other tests.
- **Predictive validity:** Good evidence that the SJT predicts subsequent performance at medical/dental school N=217, *r*=.34
  Patterson et al, in press Academic Medicine.
- **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.


0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.20
-0.25
-0.30
Managerial and Professional
Intermediate
Small employers
Low supervisory
Semi-routine
Abstract
DAT
Quant
Verbal
SJT

Widening access using SJTs

- Applicants' SES impacted their SJT scores far less than their cognitive (CAT) scores, **i.e. the SJT notably helps redress the disadvantage to lower SES applicants**
- Cohen's d ≤.20 little/no effect

Table 2 Means, standard deviations (SDs), independent-samples t-tests and effect sizes of the cognitive ability test (CAT) and situational judgement test (SJT) scores according to socio-economic status (SES)

	High SES group		Low SES group				
	n	**Mean ± SD**	***n***	**Mean ± SD**	**Mean difference**	**Cohen's *d***	**95% CI of *d***
2012 cohort							
SJT	11 966	204.76 ± 11.70	3615	203.22 ± 12.14	1.54*	0.13	0.09–0.17
CAT	11 966	645.01 ± 64.51	3615	620.34 ± 66.66	24.67*	0.38	0.34–0.42
2013 cohort							
SJT	11 756	198.82 ± 15.54	3698	195.64 ± 17.41	3.19*	0.20	0.16–0.24
CAT	11 756	677.17 ± 72.16	3698	651.51 ± 76.48	25.66*	0.35	0.31–0.39

95% CI = 95% confidence interval.
*$p < 0.01$.

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 Clinical Aptitude Test (UKCAT) in terms of three demographic variables: (i) socio-

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 the admission decision and when the selection ratio was stringent, simulated scenarios showed

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

Medical Education, 2016

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

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.

Case Study 2.

Using SJTs for selection into early careers in banking

RBS Group

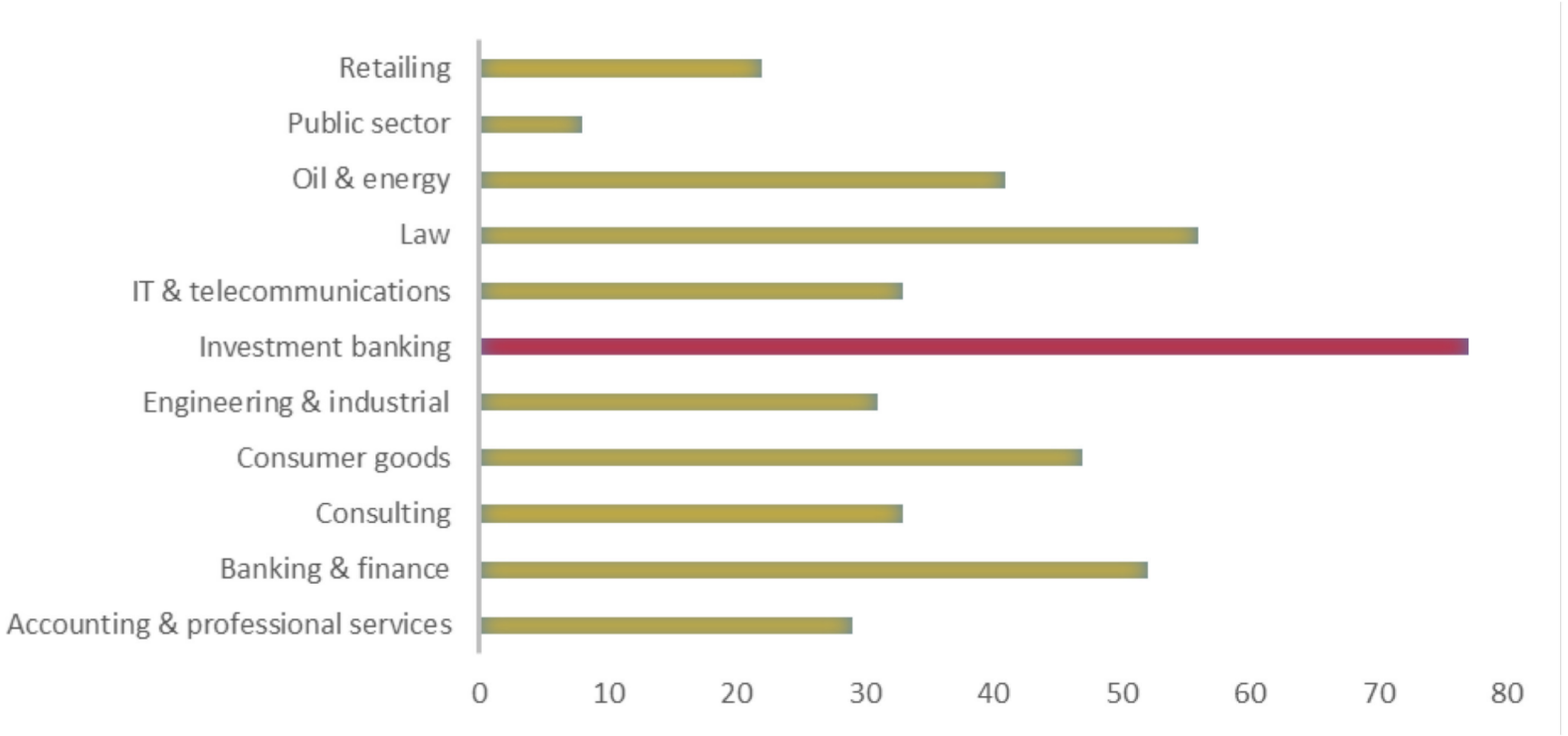
The bank you build

Work Psychology Group
Thinking differently

SES & banking sector selection

- 18% of all UK children attend a fee-paying school, in contrast to 34% of new entrants to the banking sector.
- In private equity roles, 69% of new entrants were educated privately & are from 'target' universities

Figure 3.4: Percentage of vacancies likely to be filled by graduates who had already worked for employer



Source: The Graduate Market, 2016, Highfliers[40]

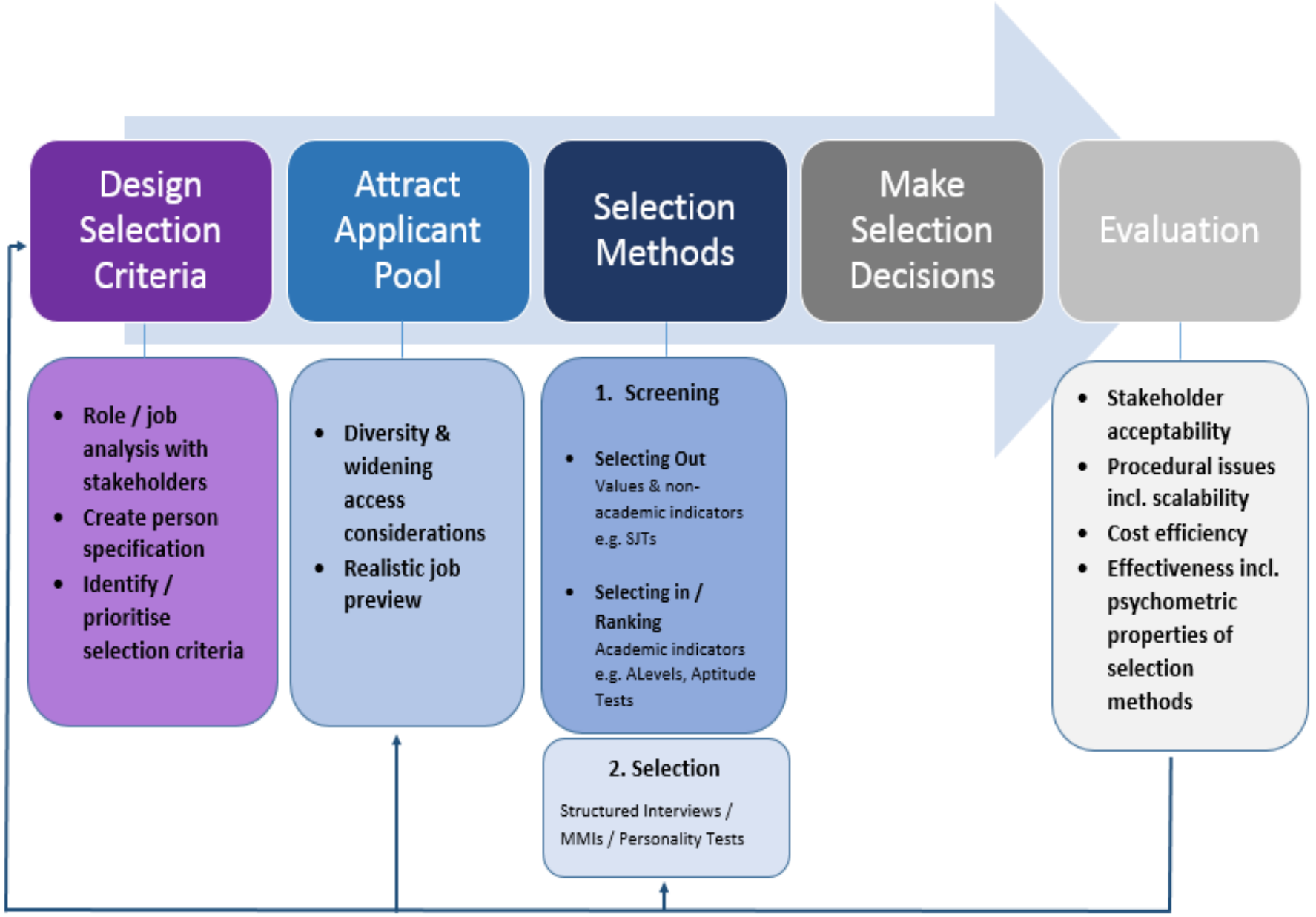
Evaluation results

- Good **psychometric properties** (the test differentiates effectively, with acceptable reliability)
- Those from **state schools (non-selective) group scored significantly higher on the SJT** than those from 'independent/private' schools ($p<.01$)
- **Females** outperform males (unlike the CAT)
- **No adverse impact** for ethnicity
- **Lower levels of candidate attrition** (i.e. greater engagement with the process & enhanced candidate experience)

What are SJTs measuring?

- SJTs measure **prosocial implicit trait policies** (**ITPs**) which are shaped by early **socialisation** (parental modelling) that teach the utility of expressing certain traits in different settings;
  - *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*
- Prosocial actions are often part of **role modelling, leadership & interpersonal exchanges** and are related to effective performance
- People with stronger ITPs about the utility of prosocial action will tend to endorse prosocial SJT response actions

A model for future design & evaluation of selection



Summary & future research

- Research on **optimal weightings & sequencing** of each method in a selection system
- Should non-academic attributes be used for '**selecting out**' & academic attributes used for '**selecting in**'?
- An SJT would need to be **heavily weighted** in selection to significantly alter the demography of those appointed
- Has the case been made more strongly in the **corporate sector?**
- Lack of evidence for use of **contextual data** in selection
- Increased focus on the role of **selection methods** in promoting **diversity** & **widening access** in recruitment

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.

Work Psychology Group
Thinking differently

Thank You

f.patterson@workpsychologygroup.com