THE UNIVERSITY of EDINBURGH

Selection & Assessment

Current Psychology 1

Tom Booth

Department of Psychology
The University of Edinburgh

Acknowledgement

> Materials for the next two weeks have been heavily influenced by the approach and materials of a Dr. David Hughes, Alliance Manchester Business School, University of Manchester.

Topics for today

- Brief introduction to me
- Introduction to organisational psychology
- Background on core ideas from individual differences
- Background on selection and assessment
- Group discussion
- Validity of selection methods
- Adverse impact in selection

Part 1: Introduction

Part 2: Key concepts from individual differences

Part 3: Selection and assessment

Part 4: Group discussion

Part 5: Validity of selection methods

Part 6: Adverse impact in selection

Me

- PhD Organisational Psychology
- Interests:
    - Org Psych
    - Health
    - Individual differences
    - Quant methods (psychometrics)

Organisational Psychology

- The scientific study of people in organisations!
- Roberts (2006) ASTMA model:
  - **A**ttraction,
  - **S**election,
  - **T**ransformation,
  - **M**anipulation, and
  - **A**ttrition

Part 1: Introduction

Part 2: Key concepts from individual differences

Part 3: Selection and assessment

Part 4: Group discussion

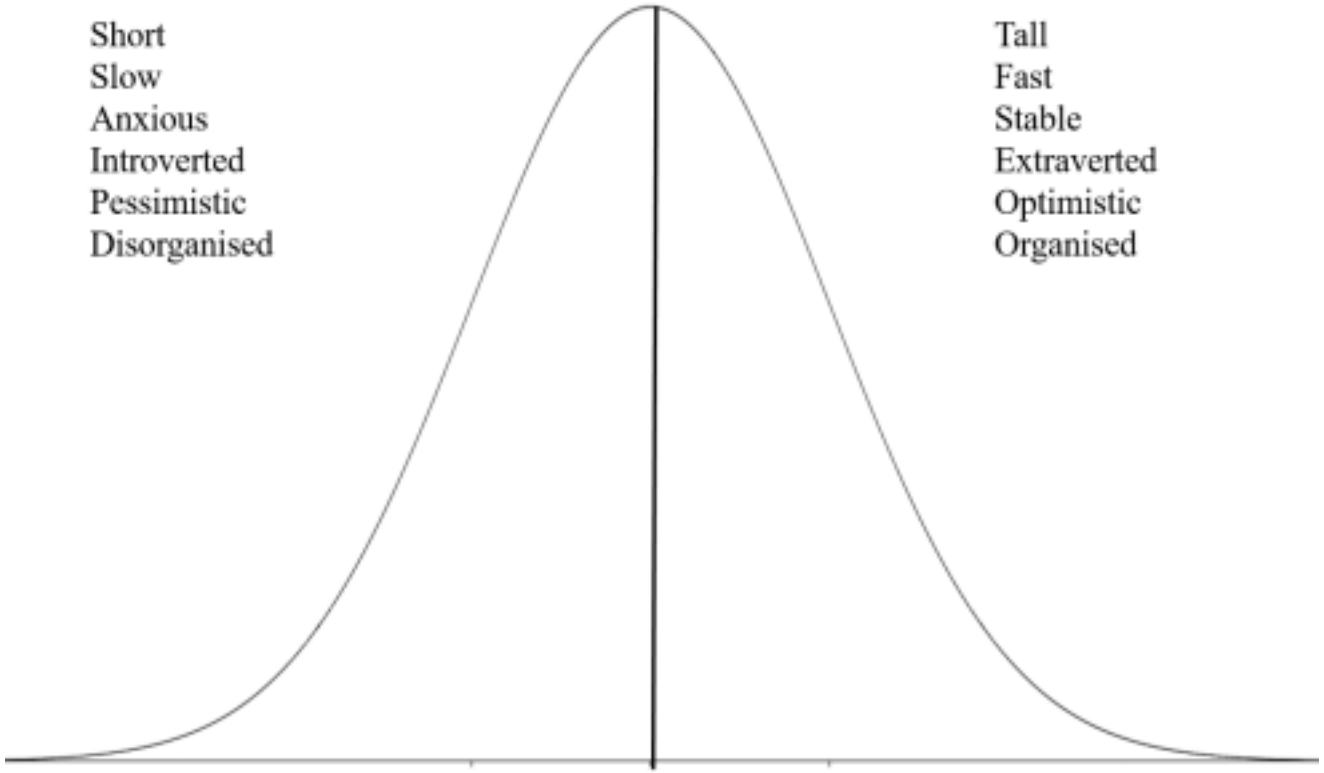
Part 5: Validity of selection methods

Part 6: Adverse impact in selection

Core principles of individual differences

- People vary!
- Understanding the dimensions and extent of individual variation in a variety of characteristics can help us explain differing outcomes
- Within "classic" experimental frameworks, individual differences are classed as error to be randomised out.
- But individual differences have been shown to be useful predictors in certain contexts.
    - Today we will look at how useful they are (or are not) in selection

Trait as a continuum



Psychometrics

- Collective term for the development and evaluation of measurement tools for individual differences.
- Often divided into tests of...
- **Typical performance** (personality, motivation, well-being)
    - Assessing how likely or frequently people may choose to behave in different ways
    - Often make use of Likert-type rating scales
- **Maximal performance** (cognitive and mental abilities, knowledge tests)
    - Attempting to assess levels of competence in a given area
    - Right/wrong answers, or speeded performance
- Wide variety of quantitative tools to analyse the resultant measures
    - Commonly factor analysis or item response theory

Personality traits

- "How shall a psychological life history be written? What processes and what structure must a full-bodied account of a personality include? How can one detect unifying threads in a life, if they exist?" **(Allport, 1968)**
- "An individual's characteristic style of behaving,thinking, and feeling"
    - **Funder (2016)** added "together with the psychological mechanisms - hidden or not - behind" it
- "Those characteristics that account for a person's consistent patterns of feeling, thinking, and behaving" **(Pervin & John, 1999)**
- "A set of individual differences that are affected by the development of an individual: values, attitudes, personal memories, social relationships, habits, and skills" **(Mischel, et al., 2004)**

Hierarchial Structure

- It is widely accepted that personality is structured hierarchically.
- At the top are broad characteristics of traits (domains)
    - These are comprised of narrower traits (facets)
    - Which are made up of items (or even narrower traits - nuances)
- Where do the traits come from?
    - responses to ratings
    - In the earliest research these were adjectives from dictionary
    - In later work, these came from the items in measures

Personlity hierarchies

- There is no single agreed upon hierarchy, but one is dominant **the Five Factor Model**

Domain	Facets
Neuroticism	Anxiety, Hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability
Extraversion	Warmth, Gregariousness, Assertiveness, Activity, Excitement-Seeking, Positive Emotions
Openness To Experience	Fantasy, Aesthetics, Feelings, Actions, Ideas, Values
Agreeableness	Trust, Straightforwardness, Altruism, Compliance, Modesty, Tender-mindedness
Conscientiousness	Competence, Order, Dutifulness, Achievement Striving, Self-Discipline, Deliberation

Cognitive abilities: Definition

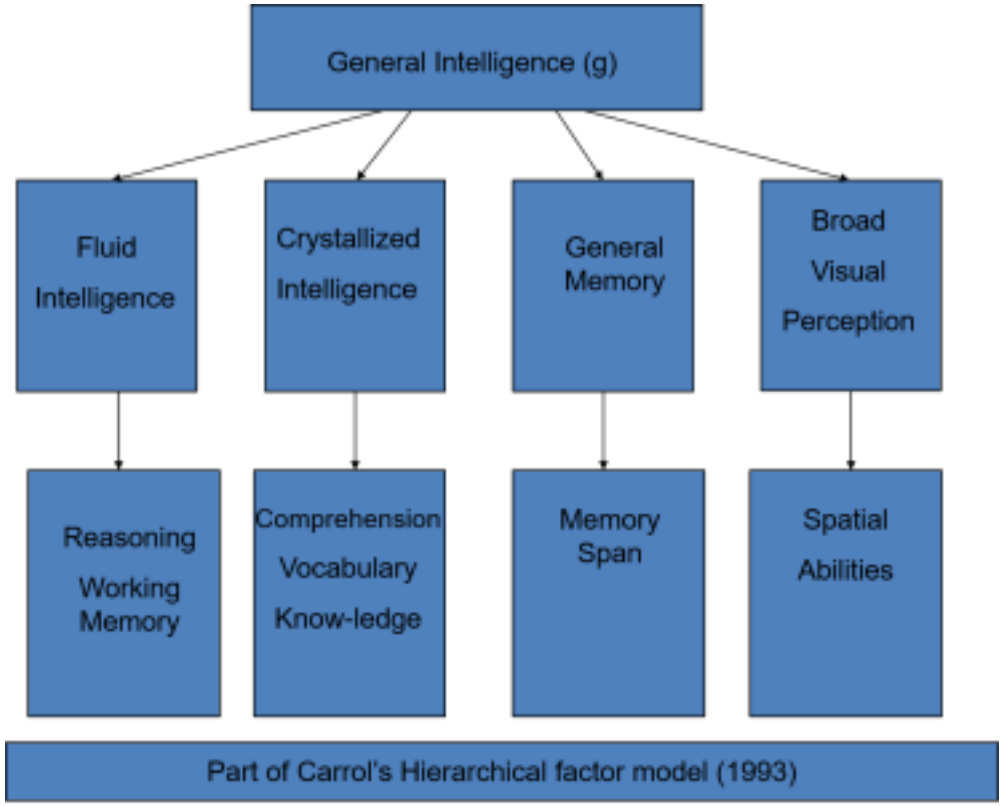
"A very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings - 'catching on', 'making sense' of things, or 'figuring out what to do.'"

(Gottfredson, 1994 consensus statement)

How is it measured?

- Vocabulary
- Relations among words
  - Similarities, analogies, opposites
- Identifying sequence progressions
  - Of numbers, letters, figural configurations
- Short-term and working memory
  - Lists of digits, unrelated words; keeping track of one thing while doing another
- Speed of very simple processing
  - Identifying or coding symbols; reaction, inspection times
- Ability to visualise transformation of shapes and figure

Organising principle (debated)



Part 1: Introduction

Part 2: Key concepts from individual differences

Part 3: Selection and assessment

Part 4: Group discussion

Part 5: Validity of selection methods

Part 6: Adverse impact in selection

What do we mean by selection and assessment?

- Finding the best candidate for the job
- We can mean different things by "best"
    - Best in terms of ability, attempting to predict job performance
    - Best in terms of 'organisational fit'
- As we will see, defining "best" is a key part of the process

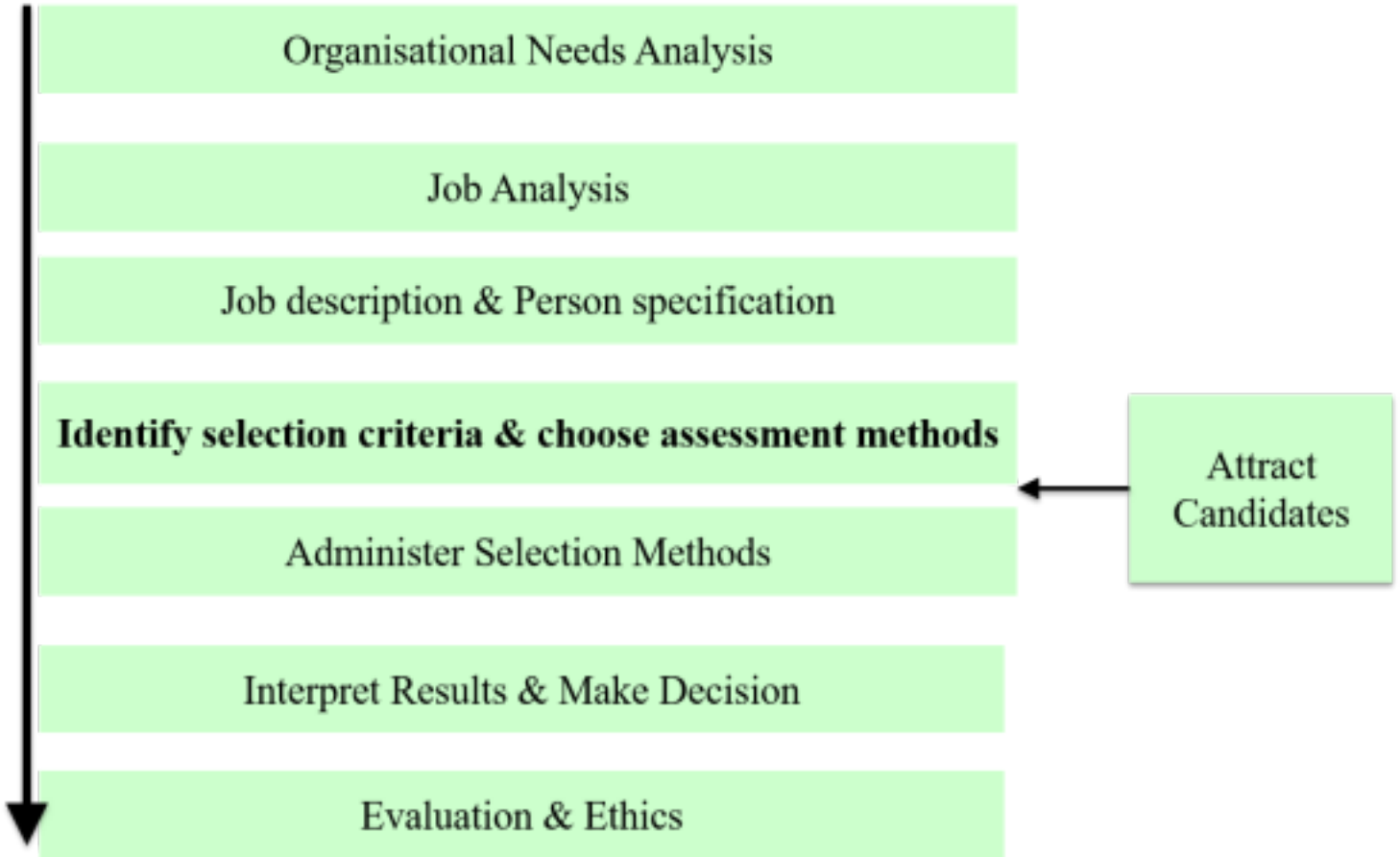
Discrimination is necessary

- But we need to be clear what we mean.
- Workplace selection is all about discriminating.
    - Identifying who is "best" requires us to discriminate candidates
- **BUT** We must ensure that we discriminate in a **FAIR** and **JUSTIFIED** manner.
    - We need to assess candidates only on job relevant skills and abilities.
- Hence we need to know what these skills and abilities are, and we need to identify assessments tools which can measure them in a fair way to all candidates.

Selection

- Process must be legally defensible
    - tools used shown to predict future job performance
    - application of the tools has been consistent to all candidates
    - can be evidenced
- Employing the 'wrong' person is costly!
    - Recruitment costs 10-40% of 1st years salary + training costs, lost productivity, knock on effect for colleagues...
    - Human level, firing people is tough
- Done effectively, good selection can save organisations money.

Selection Process



Job Analysis (1)

Purpose is to identify the skills and behaviours needed to be successful in the job

- Identifies tasks and responsibilities involved in the job
- Provides information on skills and characteristics that are important for job success
- Other factors:
    - job title,
    - hours of work,
    - direct report,
    - equipment,
    - pay & benefits etc.

Job Analysis (2)

How might we do a job analysis?

What we need to know?

- Tasks
- Skills needs
- outputs
- Performance Criteria
- Job context
- Equipment needs

How do we gather it?

- Current post holders
- Supervisors
- Expert job analysts
- Role documentation
- Observation

Methods for selecting employees

1. References
2. CV
3. Work samples
4. Job knowledge test
5. Assessment centres
6. Structured Interviews
7. Unstructured Interviews
8. Cognitive ability/ aptitude tests
9. Personality
10. Integrity test

What makes a good assessment?

- **Reliability**:
  - Precision of measurement.
- **Validity**
  - Accurate measurement (i.e., measuring what is claimed to be measured)
  - Appropriate levels of prediction (i.e., predicting future job performance)
- **Fairness**:
  - Selection methods must not unfairly discriminate against any specific subgroups
- **Candidate reactions**:
  - Often connected to something termed *face validity*
  - Selection tools might 'put off' certain candidates
- (For an organisation) Cost and other logistics

Part 1: Introduction

Part 2: Key concepts from individual differences

Part 3: Selection and assessment

Part 4: Group discussion

Part 5: Validity of selection methods

Part 6: Adverse impact in selection

Task

- Break off into groups of 4-6
    - Discuss the selection methods introduced in the last section
    - Which are the most common?
    - Identify which you think might predict performance best
    - Identify which you think might have potential issues with bias, adverse impact on candidates etc.
- Will will take 15minutes to discuss in groups, then 15 minutes as a class discussion
    - You do not have to cover all methods

Part 1: Introduction

Part 2: Key concepts from individual differences

Part 3: Selection and assessment

Part 4: Group discussion

Part 5: Validity of selection methods

Part 6: Adverse impact in selection

Evidence for personality

- Is it useful? Does it predict performance?
- **Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, (2007)**:
  - Due to the low validity and content of some items, many published self-report personality tests should probably not be used for personnel selection
- **Ones, Dilchert, Viswesvaran & Judge, (2007)**:
  - Any selection decision that does not take the key personality characteristics of job applicants into account would be deficient.

FFM and work outcomes

- **Mount & Judge (2001)**: Most comprehensive meta-analysis to-date.
- Job Performance
  - Conscientiousness: All roles (r = .23)
  - Neuroticism: All roles (r = -.13)
  - Extraversion: Management (r = .21)
  - Openness: Training (r = .33)
  - Agreeableness: Team-working (r = .34)

FFM and leadership

- **Judge, Bono, Ilies & Gerhardt (2002)**: Meta-analysed 222 correlations from 73 samples
  - Extraversion = .31
  - Conscientiousness = .28
  - Neuroticism = -.24
  - Openness to Experience = .24
  - Agreeableness = .08

Bandwidth of prediction

Correlations between personality at three levels of aggregation and Overall, Task, and Contextual job performance taken from Judge et al. (2013).

	Correlations with Job Performance								
	Overall			Task			Contextual		
Trait	Facet	Mid	Broad	Facet	Mid	Broad	Facet	Mid	Broad
Emotional Stability	.23	.12	.10	.25	.10	.08	.30	.21	.16
Extraversion	.40	.21	.19	.18	.14	.12	.49	.23	.21
Openness	.30	.10	.08	.17	.13	.12	.18	.06	.03
Agreeableness	.19	.17	.17	.24	.11	.10	.33	.18	.18
Conscientiousness	.26	.27	.26	.24	.25	.25	.33	.32	.32

Note: Facet = 6 NEO facets, Mid = DeYoung et al. (2007) factors, Broad = FFM

Findings from Judge et al (2013), Table from Hughes & Batey (2017)

Some issues

- **Social desirability**: when faced psychometrics, many people feel 'judged' and so alter their answers accordingly.
- This may happen for two main reasons:
  - **Self-deception** - individuals are overly optimistic in their self-perceptions and play down their perceived negative aspects
  - **Impression management** - individuals try to appear 'good' because they fear social disapproval.

What do others think?

Trait and rating type		r1	r2		r3
	Connelly & Ones (2010)			Oh, et al. (2011)	
Emotional Stability					
Other-rating	.14	.17	.37	.17	.24
Self-rating	.06	.11	.12	.09	.14
Extraversion					
Other-rating	.08	.11	.18	.21	.29
Self-rating	.06	.11	.12	.06	.09
Openness					
Other-rating	.18	.22	.45	.20	.29
Self-rating	.03	.04	.05	.03	.05
Agreeableness					
Other-rating	.13	.17	.31	.23	.34
Self-rating	.06	.11	.13	.07	.10
Conscientiousness					
Other-rating	.23	.29	.55	.31	.41
Self-rating	.12	.20	.23	.15	.22

Results from Connelly & Ones (2010); Oh, Wang & Mount (2011); summary table from Hughes & Batey (2015)

Cognitive ability: Predictive validity (1)

- General mental ability (GMA) is one of the best predictors of job performance (r=.51) exceeded only by work samples. (Robertson & Smith, 2001).
- In Europe the validity of GMA appears to be higher at .62. (Salgado, Anderson, Moscoso, Bertua & Fruyt, 2003).

Cognitive ability: Predictive validity (2)

Table 3. Meta-analysis results for GMA tests: occupation – job performance combinations

Occupation	*K*	*N*	*r*	*SDr*	rho	*SDrho*	%VE	90% CV	Lrho	NSD	LCV
Clerical	5	628	.14	.07	.32	.00	177	.32	.29	.09	.17
Engineer	5	542	.33	.24	.70	.42	30	.16	.64	.45	.06
Professional	4	348	.36	.18	.74	.23	61	.45	.67	.31	.28
Driver	2	293	.16	.09	.37	.00	109	.37	.34	.11	.20
Operator	9	3,105	.24	.05	.53	.00	365	.53	.48	.15	.29
Manager	5	302	.33	.01	.69	.00	200	.69	.63	.20	.37
Sales	6	483	.25	.20	.55	.31	46	.15	.50	.34	.07
Miscellaneous	7	943	.18	.08	.40	.00	166	.40	.36	.11	.22
Sample Total	43	6,644									

Bertua, Anderson & Salgado (2005)

Cognitive ability: Predictive validity (3)

Table 4. Meta-analysis results for GMA: occupation – training success combinations

Occupation	*K*	*N*	*r*	*SDr*	rho	*SD*rho	%VE	90% CV	Lrho	NSD	LCV
Clerical	8	1,989	.33	.13	.55	.11	75	.41	.50	.19	.26
Engineer	5	1,381	.39	.15	.64	.14	68	.46	.58	.23	.29
Professional	3	295	.35	.14	.59	.08	88	.49	.54	.19	.30
Driver	3	1,674	.28	.06	.47	.00	206	.47	.43	.14	.25
Operator	17	4,322	.32	.12	.54	.07	86	.45	.49	.17	.27
Skilled	12	3,086	.33	.14	.55	.14	65	.37	.50	.21	.24
Miscellaneous	14	7,258	.33	.10	.55	.00	104	.55	.50	.16	.30
Sample total	62	20,005									

Bertua, Anderson & Salgado (2005)

Cognitive ability: Predictive validity (4)

".intuition may tell [people] that personality and other non-cognitive traits are more important than GMA (Hunter & Schmidt, 1996). However... Causal analyses of the determinants of job performance show that the major effect of GMA is on the acquisition of job knowledge: People who are higher in GMA acquire more job knowledge and acquire it faster. The amount of job-related knowledge required on even less complex jobs is much greater than is generally realized. Higher levels of job knowledge lead to higher levels of job performance. Viewed negatively, not knowing what one should be doing--or even not knowing all that one should about what one should be doing--is detrimental to job performance. In addition, knowing what one should be doing and how to do it depends strongly on GMA."

(Schmidt & Hunter, 2004)

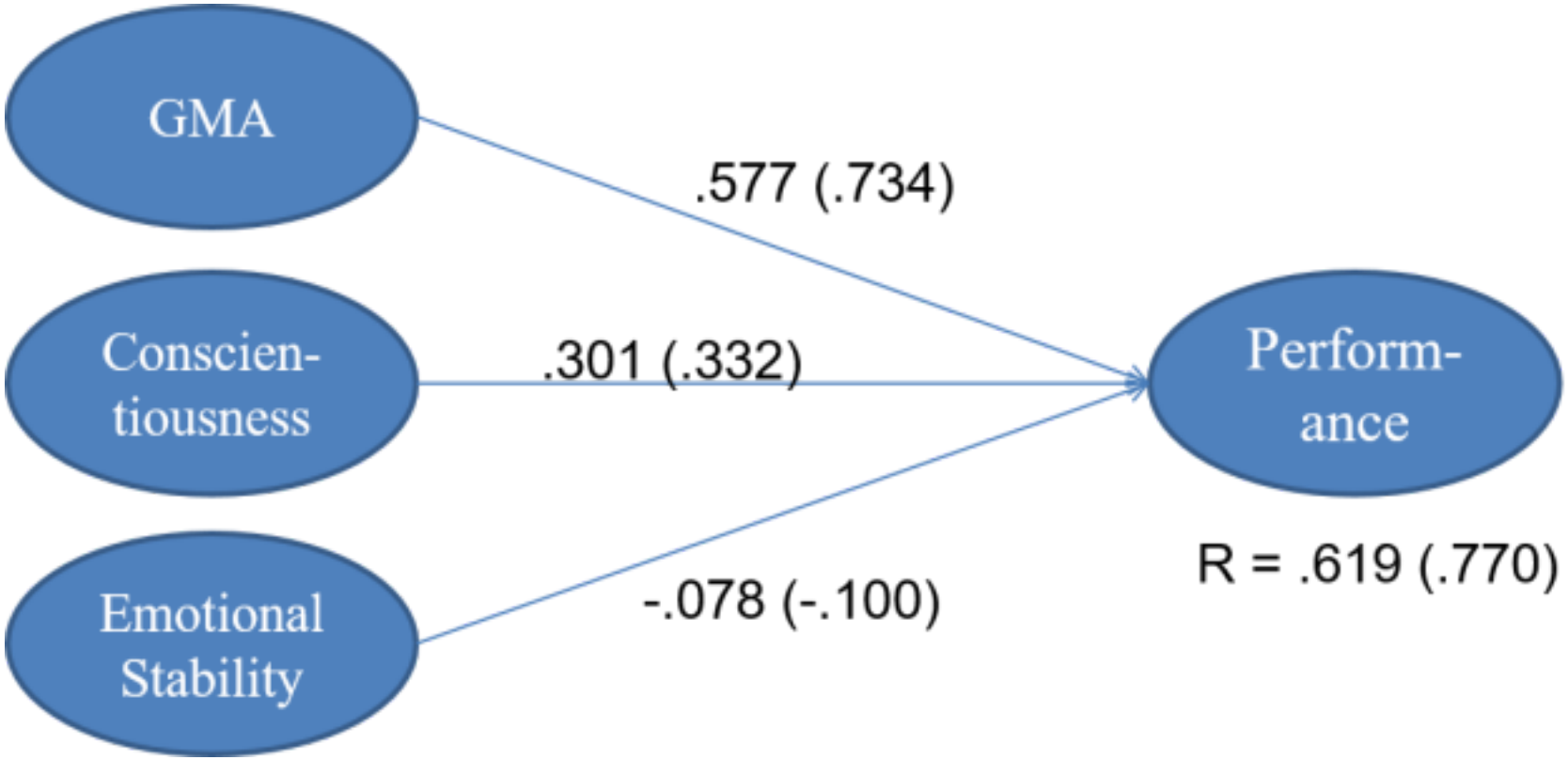
Big picture?

Selection method	r
Work samples	.54
General Mental Ability	.51
Structured Interviews	.51
Job knowledge	.48
Personality	.40
Assessment centres	.37
Unstructured Interviews	.38
Biographical data	.35
References	.26
Graphology	.02

Selection method combinations	r
General Mental Ability + Personality	.64
General Mental Ability + Structured Interviews	.63
General Mental Ability + Work Samples	.60

From Schmidt & Hunter (1998)

Schmidt, Hunter & Oh (2008)



- 1,577 studies
- corrections for host of statistical artefacts
- Personality adds 5-8%

Arnold et al., (2005)

Selection method	Evidence for criterion-related validity	Applicant reactions	Extent of use
Structured interviews	High	Moderate to positive	High
Cognitive ability	High	Negative to moderate	Moderate
Personality tests	Moderate	Negative to moderate	Moderate
Biodata	Can be high	Moderate	Moderate
Work sample tests	High	Positive	Low
Assessment centres	Can be high	Positive	Moderate
Handwriting	Low	Negative to moderate	Low
References	Low	Positive	High

Part 1: Introduction

Part 2: Key concepts from individual differences

Part 3: Selection and assessment

Part 4: Group discussion

Part 5: Validity of selection methods

Part 6: Adverse impact in selection

What do we mean by adverse impact (AI)?

Broadly speaking, adverse impact has been defined as subgroup differences in selection rates (e.g., hiring, licensure and certification, college admissions) that disadvantage subgroups protected under Title VII of the 1964 Civil Rights Act. Protected subgroups are defined on the basis of a number of demographics, including race, sex, age, religion, and national origin (Uniform Guidelines, 1978).

(Outtz & Newman,)

Formal definition of AI

$$AI = \frac{SR_1}{SR_2}$$

- AI = adverse impact
- SR_1 = selection ratio for minority group
- SR_2 = selection ratio for majority group

$$SR = \frac{\text{Number of people hired from a group}}{\text{Total number of applicants from group}}$$

- AI values of 1.0 are optimal (identical selection ratio)
  - In the US, AI of 0.80 is used as a marker when considering compliance (Roth, Bobko, & Switzer, 2006)

Why is this considered a major issue

- The definitions concern selection rates, so we may ask why this is a problem
- It becomes a potential issue when we also note that many of the tools used for selection (particularly cognitive tests), have been shown to have mean score level differences across groups.
    - The magnitude of these differences is much debated
    - As are the reason why such differences might be observed

Outtz & Newman

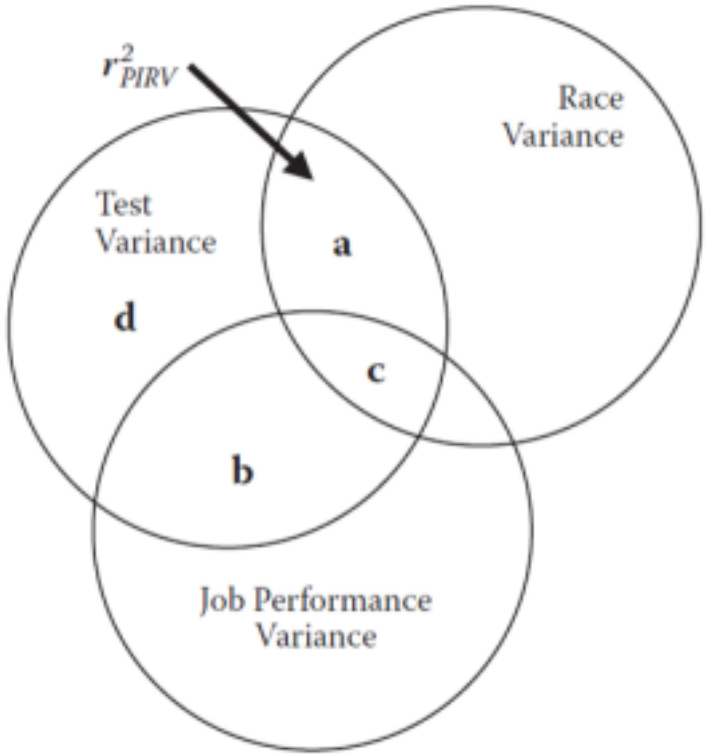


FIGURE 3.1
Performance-irrelevant race-related variance in test scores.

Theory of AI (Outtz & Newman)

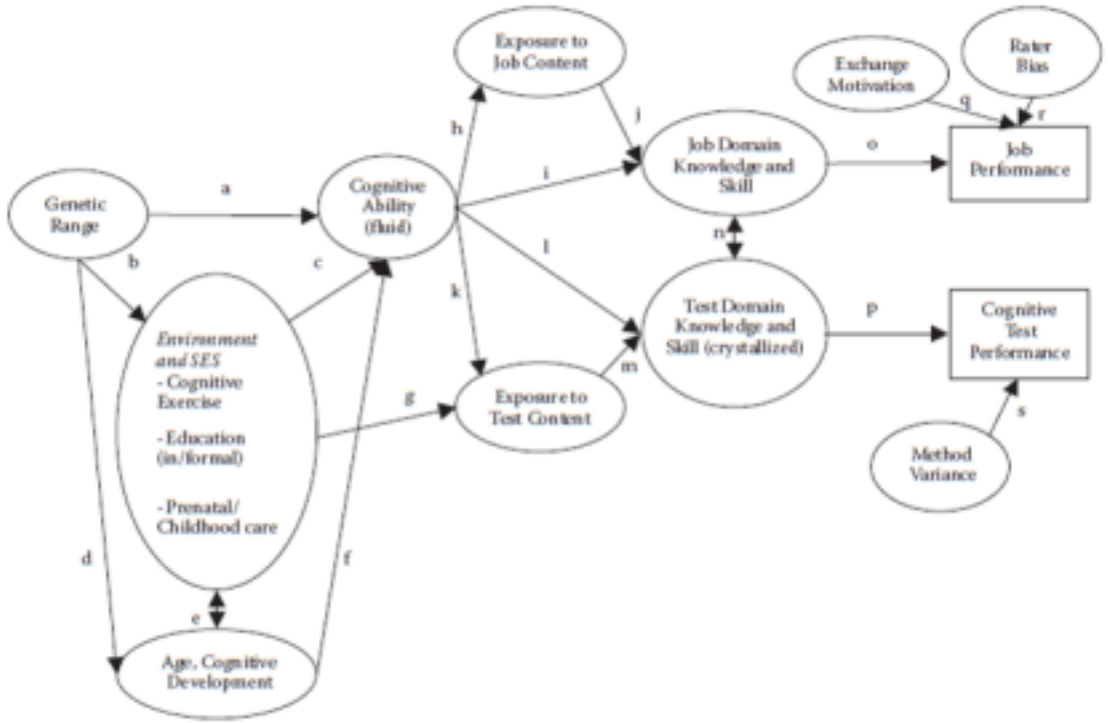


FIGURE 3.5
Second-generation model of adverse impact (without theory of race).

Reducing adverse impact

- Suggestions range from..
  - the simple - use multiple measures
  - to the extraordinarily complex major societal interventions in key areas around education and opportunity
- Outtz & Newman argue that r_{PRIV} should be a focal parameter in psychometric measure development for selection.
  - The closer a measure is to performance on the job, the better, as this should reduce r_{PRIV} by making more test variance job related
- Aguinis & Smith (2007): Use data on test bias explicitly in the decision making process.

Summary

- Today we have outlined:
  - Background to individual differences in context of selection
  - Discussed a variety of selection tools
  - Compared the predictive validity of these tools
  - Defined adverse impact
  - Discussed adverse impact with respect to various selection tools

That's all for this week