

Swift Occupational Ability

Course Workbook

training[•]

training[•]

Introduction: Saville Assessment	1
Section 1: Introduction to Testing	5
Section 2: Job Analysis	13
Section 3: Assessment Choice	23
Section 4: Test Administration	45
Section 5: Test Scores	55
Section 6: Test Feedback	61
Section 7: Test Norms	65
Section 8: Correlation	85
Section 9: Reliability	93
Section 10: Validity	107
Section 11: Approaches to Online Testing	129
Section 12: Best Practice & Ethics	135
Section 13: Next Steps	143
Section 14: Post-course Work	147
Section 15: Review Worksheets	207
Section 16: Resources	227

Introduction: Saville Assessment

Objectives

By the end of the course, you will be able to:

- Choose assessments
- Administer tests
- Interpret test scores
- Feedback test results
- Apply tests fairly
- Apply to the British Psychological Society for external certification (optional)
- Gain access to our extensive Aptitude suite and Work Strengths tool





*To dive deeper than
humanly possible
and bring talent to
the surface.*

About Saville Assessment, A Willis Towers Watson Company

- 2004: Founded by Professor Peter Saville
- 2005: Wave launched
- 2007: Swift combination tests launched
- 2009: Item-banked aptitude tests
- 2013: Situational Judgement Tests launched
- 2015: New tests, new technology
- 2015: A Towers Watson Company
- 2016: Willis Towers Watson

Course Overview

Day 1

Introduction to Testing



Job Analysis



Assessment Choice



Test Administration



Test Administration Overview



Test Scores



Test Feedback

Day 2

Test Norms



Feedback
Practical Session



Correlation



Reliability



Validity



Approaches to Online Testing



Best Practice & Ethics



End of Course
Assessment

Notes

Eligibility for BPS Certification

Following successful completion of this course and submission of the post-course work within the stipulated time limits, delegates are able to apply to the British Psychological Society (BPS) for Test User: Occupational Ability certification. Please note that the BPS has separate additional charges for this process which are not covered in Saville Assessment course fees. Please visit <https://ptc.bps.org.uk/> for further details of current charges.

About Saville Assessment

The journey of Saville Assessment started in 2004 when a team of assessment specialists came together. The team comprised experts in Occupational Psychology, Business Consulting and Information Technology, with the goal of transforming and revolutionising assessment around the world.

Our assessment tools are available in over 40 languages; please contact us for more information.

In 2015, Saville Assessment was acquired by Towers Watson and now forms part of the Willis Towers Watson global organisation.

A Brief History

2004 - Saville Consulting is founded

'Assessment Guru' Professor Peter Saville recruited a team of assessment experts/psychometricians to deliver his vision of transforming assessment around the world.

2005 - Wave

A new era of personality questionnaires arrives, offering the highest validity on the market and the deepest insight into an individual's motives, talents and workplace potential.

2007 - Swift combination aptitude tests

Faster, smarter ability testing boasting a fresh, modern look and feel, and the only portfolio to include combination tests measuring several sub-areas in one assessment.

2009 - Item-banked aptitude tests

Introduction of item-banks across our aptitude test portfolio to ensure greater security in online assessment.

2013 - Situational Judgement Tests

Bespoke, multi-media SJTs combining psychometric expertise with the latest technology breaks boundaries with a fast, engaging, powerfully branded volume assessment tool.

2015 - New tests, new technology

The first psychometric test publisher to have tablet-administered assessments and lead the way with utilising technology.

2015 - 2017 - Saville Assessment, A Willis Towers Watson Company

Became the talent assessment part of the leading global advisory, broking and solutions company, helping clients around the world turn risk into a path for growth.

Section 1: Introduction to Testing

Talent Trends and Challenges

- Talent is global
- Applicant numbers per role are increasing
- Organisations are assessing behaviours, cultural fit, values
- Emphasis on employer brand
- Candidate experience is critical
- Recruitment processes are speeding up
- Online, mobile and remote assessment is the norm
- Security of assessment materials is still a risk
- Social media has implications
- Selection data is increasingly used for onboarding or employability
- Responsibility for development is devolved to line managers
- Increasing expectation that career management is self-led
- The 'gig' economy is emerging
- Everyone's talking about Big Data

Projective Tests – Inkblot Test



Projective Tests – Thematic Apperception Test



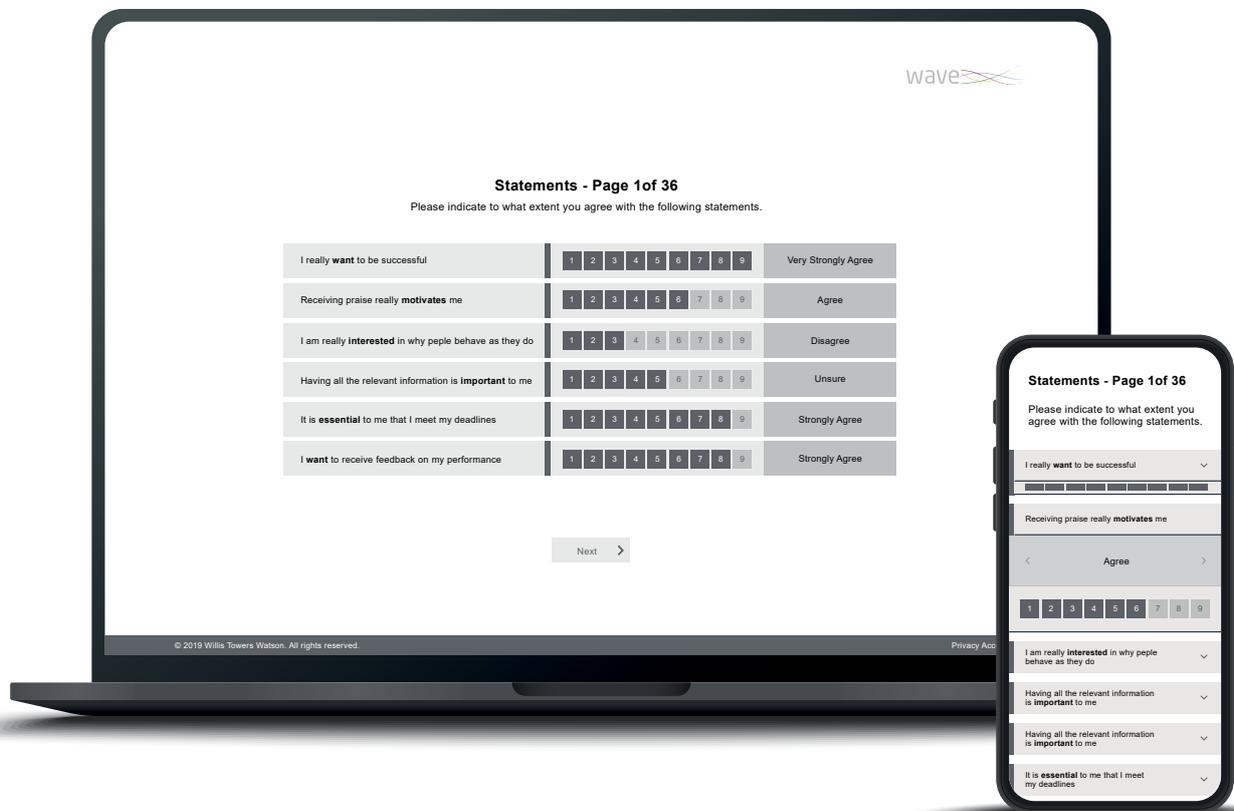
What is a Psychometric Test?

An assessment of a psychological attribute, typically scored using a numerical scale or category system, to describe individual differences.

'Will Do' Assessments of Typical Performance

- Include self-report questionnaires without time limit
- 'Right' and 'wrong' can vary depending on context

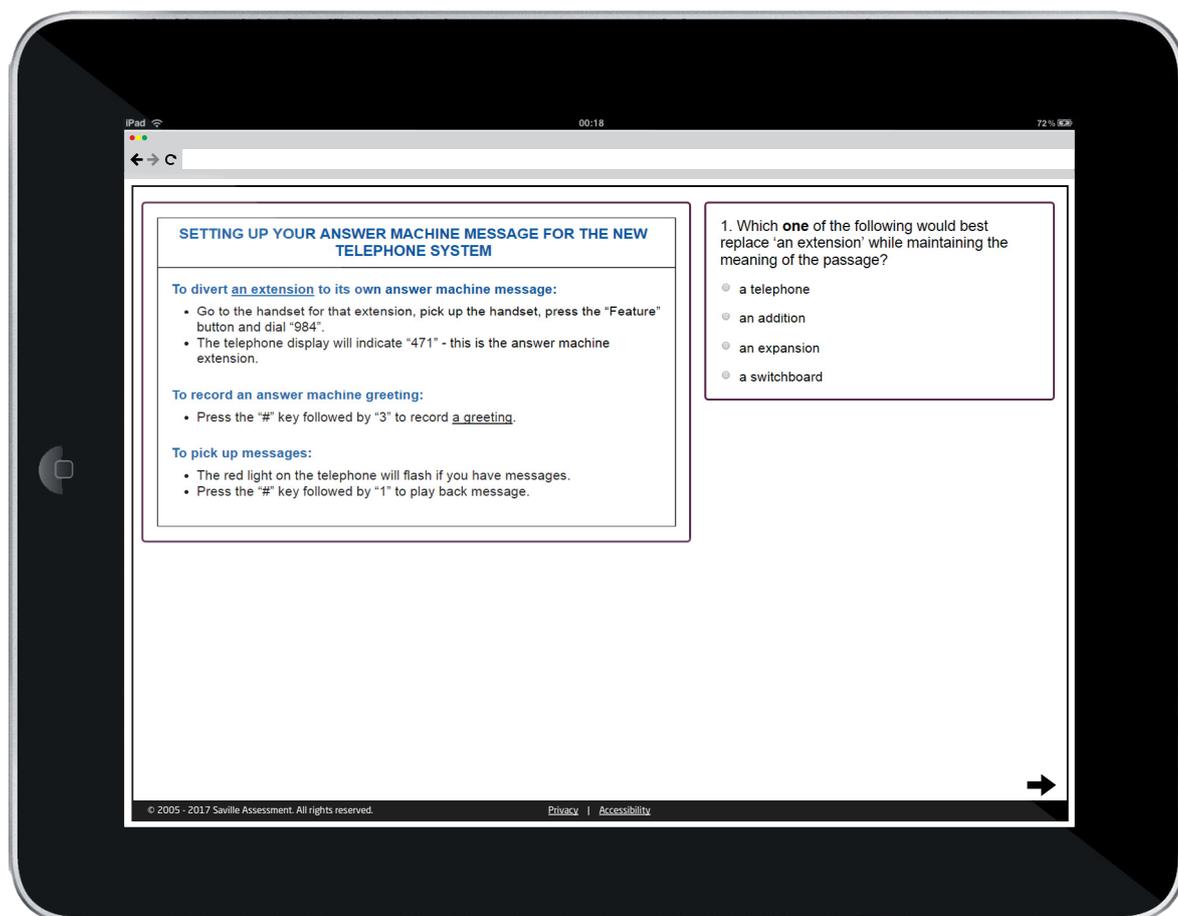
Wave Professional Styles example:



'Can Do' Ability Tests of Maximum Performance

- Include ability tests of aptitude, IQ and attainment:
 - Aptitude: predict what someone will be able to learn or do in the future, e.g. Saville tests
 - IQ: current level of intellect/cognitive ability, e.g. Wechsler Adult Intelligence Scale
 - Attainment: measure current level of knowledge understanding or skill, e.g. driving test
- Often with strict time limits
- Clear right or wrong answers

Verbal analysis example:



Key Facts: Aptitude Tests

- Benchmarks against external group
- Single most valid predictor of work performance
- Measures lots of different types of ability
- Efficient online assessment
- Fair and consistent treatment of candidates
- Supplements other sources of information
- Sophisticated question banking to counter cheating

Notes:

Notes

Methods of Assessment

Graphology is handwriting analysis; this method previously enjoyed some popularity in France but is rarely used in practice.

Interviews can take several forms. The most common are listed below:

- **Biographical interviews** typically ask questions around a CV or application form, covering areas such as education, work experience, interests, etc.
- **Behavioural event (competency) interviews** ask candidates to provide examples of situations where they displayed particularly effective behaviours. Normally, this is based on a structured interview guide and a reasonably structured scoring mechanism based on the evaluation of the quality of the answers.
- **Situational interviews** put the candidate in particular situations that they are likely to find themselves in for the role and they are asked to role play what they would do. The scoring can be extremely structured, with candidates being awarded points for matching part of the 'perfect answer'.

Projective tests such as the Rorschach and the Thematic Apperception Test give candidates stimuli (pictures) that are open to interpretation. Candidates provide their interpretation of the picture, which is evaluated by the interviewer/assessor; it is claimed that this method uncovers individuals' unconscious needs or drives.

Biographical data includes information such as education, experience, training, etc. which can be collected via CV or structured application form.

Aptitude tests look at different reasoning abilities, e.g. verbal, numerical, diagrammatic.

Interviews

Advantages

Interviews provide the opportunity to get precise answers to questions and iron out any inconsistencies in other assessment data on the individuals. As well as this detailed evaluation and challenge of the candidates, interviews are particularly good at uncovering more people-oriented social competencies such as 'Articulate' and 'Engaging'

as well as picking up on verbal and non-verbal cues that candidates can give away when discussing particular topics (e.g. discomfort, lack of enthusiasm when talking about networking).

Interviews create a social experience and give the company a human face with which a candidate can identify. Clearly, how positive or negative this social experience is can impact on acceptance rates following job offers. It also allows candidates to ask questions of their potential employers that may be critical when it comes to deciding whether this is the best employer for them.

Interviews are also an important starting point for the effective establishment of the 'psychological contract' between employee and employer, which helps sustain employee motivation over time.

Disadvantages

Most of the time, interviewers are human and different humans care about different things. As a result, interviewers tend to have different views on which attributes are most important for a particular job. Undoubtedly, a good specification of the criteria that relate to success for a particular post is a great help in this regard, and better still is an interview guide that standardises the key questions asked. However, even with this, interviewers tend to weight the same information somewhat differently when it comes to evaluation and scoring.

A common problem is the overweighting of academic qualifications and previous work experience. Academic qualifications vary from having little to no link with performance in the job. Experience may be important, however fine judgements are often made here that seem valid on the face of it but actually make little or no difference to how the candidate will perform in their new role.

Another common problem is that interviewers can ask closed or leading questions, confirming their own stereotypes or preconceptions about the candidate. Interviewers typically talk too much and both what they say and the questions they ask tend to influence the responses of candidates.

Interviewers tend to prefer people they perceive to be similar to themselves and are likely to make decisions on a candidate's suitability after only a few minutes.

Video Interviewing

Interest in video interviewing has grown steadily over the past few years, particularly in the volume recruitment space and the trend is set to continue. Video interviewing brings with it its own benefits and challenges. Candidates going through video interviews perceive those organisations using the process to be innovative and modern. Where the process is mobile or tablet enabled, interviews are particularly easy for individuals to access and complete. The same goes for assessors, who can access candidate videos from anywhere and at any time, provided they have internet access and sound. Assessors often feel like they can make fairer assessments because evidence is captured more thoroughly and videos can be paused or rewound when necessary. There is also less of a logistical challenge in coordinating candidate and assessor diaries.

There are, of course, always technical issues to arise from such processes, with candidates sometimes needing to be re-interviewed, which may present an unfair advantage at times. Whilst some candidates perceive video interviewing to be modern and innovative, others are put off by this medium and so organisations can see high dropout rates here.

For more information on Saville Assessment's Interview and Assessor Skills training, call 020 8619 9000.

Test Types

Will Do Tests

These measure typical performance, examples of which are listed below:

- **Interest inventories/questionnaires** measure the things an individual is interested in. This type of information may be useful in career guidance
 - **Personality questionnaires** look at styles of behaviour, for example the Occupational Personality Questionnaire (Saville P, et al, 1984) and the Professional Styles and Focus Styles versions of Saville Assessment's Wave.
 - **Motivation questionnaires** measure what people want to do. Note: this can also be measured by the Wave questionnaire detailed above.
 - **Rating scales** look at measures of job performance.
 - **Attitude surveys** are often of great interest in market research.
- **360 degree assessments** ask for ratings from bosses, colleagues and subordinates. Saville Assessment has developed the Wave Performance 360 questionnaire to gather self and other ratings online.

Can Do Tests

These assess maximum candidate performance, examples of which are listed below:

- **Aptitude tests** measure abilities that underpin future potential – examples include Saville Assessment's verbal, numerical and diagrammatic analysis tests.
- **Achievement/attainment tests** look at an individual's level of current knowledge.
- **Intelligence tests (IQ)** are a mixture of aptitude and attainment.
- **In-tray exercises/business simulation exercises** are tests which assess skills at particular tasks and are often very useful in assessment centres.
- **Work sample tests** present applicants for a job with a sample of the work they will be expected to undertake in the job.
- **Trainability tests** assess how well individuals respond to training.

Section 2: Job Analysis

Job Analysis

- Job analysis is a multi-method approach that is used for different purposes including:
 - Defining role profiles/job descriptions/person specifications
 - Job sizing for pay grading
 - Developing a framework of criteria for assessment e.g. behavioural competencies
- In assessment, good job analysis focuses on things that can be defined clearly and measured well

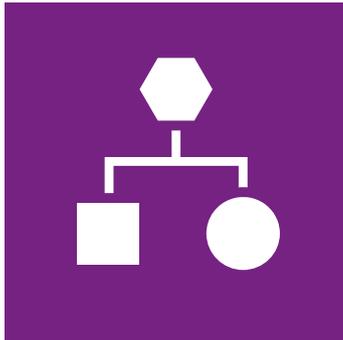
Common Methods of Job Analysis

- Structured interviews
 - Job holders e.g. critical incident identification
 - Line managers e.g. repertory grid comparisons
- Job content reviews
 - Diaries
 - Observing the job
 - Doing the job
 - Task/job analysis questionnaires
- Validation research

Notes:

Wave Job Profiler

- Online tool (15 minutes)
- Survey different stakeholders
- 7-point 'Importance' rating
- Can add comments



Job Role Profile - Solving Problems							
Solving Problems	Important			Important			
	Not	Marginally	Fairly	Important	Very	Extremely	Critically
Evaluating Problems	Examining Information Processing Information; Asking Probing Questions; Finding Solutions						
				5	6	8	
Evaluating Problems	Documenting Facts Writing Fluently; Understanding Logical Arguments; Finding Facts						
				7	7	9	
Evaluating Problems	Interpreting Data Quantifying Issues; Applying Technology; Evaluating Information Objectively						
				7	7	8	
Investigating Issues	Developing Expertise Taking Up Learning Opportunities; Acquiring Knowledge and Skills; Updating Specialist Knowledge						
				4	6	7	
Investigating Issues	Adopting Practical Approaches Applying Practical Skills; Learning by Doing; Applying Common Sense						
				3	4	5	
Investigating Issues	Providing Insights Continuously Improving Things; Identifying Key Issues; Making Intuitive Judgements						
				4	6	7	
Creating Innovation	Generating Ideas Producing Ideas; Inventing Approaches; Adopting Radical Solutions						
				3	5	7	
Creating Innovation	Exploring Possibilities Developing Concepts; Applying Theories; Identifying Underlying Principles						
				4	6	6	
Creating Innovation	Developing Strategies Forming Strategies; Anticipating Trends; Envisaging the Future						
				6	7	10	

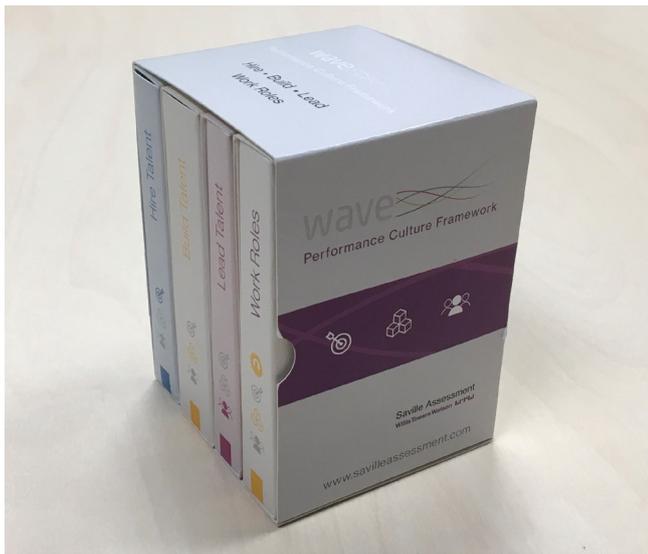
○ Boss / Job Holder ◇ Stakeholder ◻ Report

Job Role Profile - Reasoning at Work							
Reasoning at Work	Important			Important			
	Not	Marginally	Fairly	Important	Very	Extremely	Critically
Working with Information	Working with Words Understanding Word Meaning; Comprehending Text; Making Verbal Inferences; Evaluating Written Materials; Comparing Arguments						
					6	7	8
Working with Information	Working with Numbers Understanding Tables; Comprehending Graphs; Making Numerical Inferences; Evaluating Quantities; Comparing Data						
					6	9	
Working with Information	Working with Details Checking Letters and Text; Checking Numbers and Tables; Checking Codes and Symbols; Identifying Mistakes; Classifying Information						
				5	7	7	
Working with Information	Working with Systems Understanding Logical Rules; Comprehending Process Diagrams; Identifying Causes; Finding Faults; Comparing Flowchart Sequences						
					7	7	9
Working with Things	Working with Designs Estimating Lengths and Angles; Recognising Rotated Shapes; Visualising Three-Dimensional Objects; Inspecting Objects; Designing Things						
				4	4	5	
Working with Things	Working with Equipment Understanding Mechanical Problems; Comprehending Physical Principles; Estimating Movement of Objects; Using Tools; Operating Machinery						
				4	4	4	

○ Boss / Job Holder ◇ Stakeholder ◻ Report

Wave Card Deck

<p>Adjusting to Change</p> <ol style="list-style-type: none"> Thinking Positively Embracing Change Inviting Feedback <p>8</p>	<p>Building Relationships</p> <ol style="list-style-type: none"> Interacting with People Establishing Rapport Impressing People <p>4</p>	<p>Driving Success</p> <ol style="list-style-type: none"> Taking Action Seizing Opportunities Pursuing Goals <p>12</p>	<p>Creating Innovation</p> <ol style="list-style-type: none"> Generating Ideas Exploring Possibilities Developing Strategies <p>3</p>
<p>1 - Not Important</p> <p>23</p>	<p>4 - Important</p> <p>26</p>	<p>7 - Critically Important</p> <p>29</p>	<p>Hire Talent Question Card</p> <p>How important are these areas in the work role?</p> <p>19</p>



Case Study: Job Analysis

You are required to design an assessment process for one of the following vacancies.

- Call Centre Customer Support Assistant
- Senior Web Developer

Full job descriptions can be found on the following pages. Before you design your process, you'll need to do some job analysis. Normally, you would conduct job analysis using a number of different methods involving a number of different stakeholders. Card sorts are a useful way of quickly gathering opinions from individuals or groups. Have a go at one now yourself to design your person specification. Use the steps listed to help you.

1. Review your job description
2. Use the Wave card deck to identify up to seven key competencies (four behaviours and three abilities)
3. List your key competencies in the space below

Job Title: _____

Key Competencies:

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.

Performance Culture Framework Card Deck

The cards are organized as follows:

- Behavioral Section 1:** Solving Problems
- Behavioral Section 4:** Influencing People
- Behavioral Section 7:** Adapting Approaches
- Behavioral Section 10:** Delivering Results
- Skill Section 1:** Evaluating Problems
 - 1 Examining Information
 - 2 Documenting Facts
 - 3 Interpreting Data
- Skill Section 4:** Building Relationships
 - 1 Interacting with People
 - 2 Establishing Rapport
 - 3 Impressing People
- Skill Section 7:** Showing Resilience
 - 1 Conveying Self-Confidence
 - 2 Showing Composure
 - 3 Resolving Conflict
- Skill Section 10:** Processing Details
 - 1 Meeting Timescales
 - 2 Checking Things
 - 3 Following Procedures
- Skill Section 2:** Investigating Issues
 - 1 Developing Expertise
 - 2 Adopting Practical Approaches
 - 3 Providing Insights
- Skill Section 5:** Communicating Information
 - 1 Convincing People
 - 2 Articulating Information
 - 3 Challenging Ideas
- Skill Section 8:** Adjusting to Change
 - 1 Thinking Positively
 - 2 Embracing Change
 - 3 Inviting Feedback
- Skill Section 11:** Structuring Tasks
 - 1 Managing Tasks
 - 2 Upholding Standards
 - 3 Producing Output
- Skill Section 3:** Creating Innovation
 - 1 Generating Ideas
 - 2 Exploring Possibilities
 - 3 Developing Strategies
- Skill Section 6:** Providing Leadership
 - 1 Making Decisions
 - 2 Directing People
 - 3 Empowering Individuals
- Skill Section 9:** Giving Support
 - 1 Understanding People
 - 2 Team Working
 - 3 Valuing Individuals
- Skill Section 12:** Driving Success
 - 1 Taking Action
 - 2 Seizing Opportunities
 - 3 Pursuing Goals

Ability Section Cards

13

swift 
Ability Dimension
 Working with Information



©2019 Saville Assessment, Willis Towers Watson. All rights reserved

16

swift 
Ability Dimension
 Working with Things



©2019 Saville Assessment, Willis Towers Watson. All rights reserved

Working with Words
 Verbal Aptitude

A	Understanding Word Meaning
B	Comprehending Text
C	Making Verbal Inferences
D	Evaluating Written Materials
E	Comparing Arguments

13

Working with Systems/Logic
 Diagrammatic Aptitude
 Abstract Aptitude

A	Understanding Logical Rules/Sequences
B	Comprehending Process Diagrams/Processes
C	Identifying Causes/Rules
D	Finding Faults
E	Comparing Flowchart Sequences

16

Working with Numbers
 Numerical Aptitude

A	Understanding Tables
B	Comprehending Graphs
C	Making Numerical Inferences
D	Evaluating Quantities
E	Comparing Data

14

Working with Designs
 Spatial Aptitude

A	Estimating Lengths and Angles
B	Recognizing Rotated Shapes
C	Visualizing 3D Objects
D	Inspecting Objects
E	Designing Things

17

Working with Details
 Error Checking Aptitude

A	Checking Letters and Text
B	Checking Numbers and Tables
C	Checking Codes and Symbols
D	Identifying Mistakes
E	Classifying Information

15

Working with Equipment
 Mechanical Aptitude

A	Understanding Mechanical Problems
B	Comprehending Physical Principles
C	Estimating Movement of Objects
D	Using Tools
E	Operating Machinery

18

Job Description: Customer Support Assistant

Role: Call Centre Customer Support Assistant, £17,000

Location: Manchester, UK

Report to: Customer Support Team Leader

Job Description:

Leading insurance company IzaQ are seeking a highly motivated individual to join their customer services team. The team are responsible for supporting existing customers with changes to their accounts and policies, initial claim enquiries and complaints.

You must have exceptional customer services skills, including an excellent telephone manner and strong attention to detail. You will be working in a fast-paced call centre environment and must be able to handle high volumes of enquiries in a polite and efficient manner. At IzaQ, we put our customers first and pride ourselves on delivering outstanding customer services. We are seeking Customer Support Assistants who share our passion.

Key Responsibilities:

- Answering incoming customer calls and emails, responding in a timely and polite manner
- Handling customer enquiries and providing information on our products and services
- Managing and resolving customer complaints
- Entering and updating customer information accurately on the database
- Processing customer queries using our online system
- Providing instant quotes to prospective customers
- Identifying and escalating priority issues
- Documenting call information
- Providing a consistently excellent service to our customers

Required Skills and Experience:

- Minimum 5 GCSEs or equivalent
- Fluent in English
- Proficient in using computer systems
- Excellent verbal and written communication skills
- Strong numerical abilities
- Strong attention to detail
- Problem solving skills
- Able to work quickly and accurately

Desirable Skills:

- Previous customer services experience
- Experience of working in a call centre environment
- Knowledge of insurance services

Notes:

Job Description: Senior Web Developer

Role: Senior Web Developer, £52,000

Location: London, UK

Report to: Web Manager

Job Description:

Highware Inc. is looking for a motivated Senior Web Developer to lead one of our development teams in assisting clients with their online needs. As a Senior Web Developer at Highware Inc., your team will design, build and maintain websites and website applications for a broad client base. Your team will work for a variety of businesses and every day will be different: you could create a secure online shopping website one day and set up a company intranet for staff the next day.

Highware Inc. is looking for an experienced Senior Web Developer with excellent web and database programming skills, and a good appreciation of design, usability and interactivity. We are looking for someone creative who can turn client ideas into workable plans and offer solutions to complex client requirements and issues. You will have to be willing to work flexibly but in an organised manner and will have to be able to follow exact instructions from written client briefs.

Key Responsibilities:

- Developing and delivering effective solutions for clients
- Ensuring clients' websites are integrated smoothly with existing networks
- Managing a small team of Junior Developers to meet multiple client deadlines
- Dealing with user access and security
- Understanding the root cause of any technical issues and applying appropriate fixes
- Communicating with clients on project issues
- Producing monthly billing reports for the Management Team and managing project budgets
- Analysing and reporting on solution effectiveness

Required Skills and Experience:

- Degree in an IT-related subject
- Strong written & verbal communication skills
- The ability to work with deadlines
- Excellent HTML & CSS coding skills
- Experience with JavaScript
- Adobe Photoshop

Desirable Skills:

- Interested in keeping up to date with advances in computer technology
- An appreciation of commercial pressures
- Previous managerial experience

Notes:

Notes

Job Analysis

Job Analysis is a detailed process to identify and determine the particular job duties and requirements, and the relative importance of these duties for a given job. An important concept of Job Analysis is that the analysis is conducted on the job not the person. While data may be collected from incumbents through interviews or questionnaires, the product of the analysis is a description or specification of the job, not a description of the person.

Job Analysis is an essential pre-requisite to choosing which psychometric tests and questionnaires to use. There are several ways to conduct a job analysis: interviews with incumbents and supervisors, questionnaires (structured, open-ended, or both), observation, and gathering background information such as duty statements or classification specifications. In job analysis conducted by HR professionals, it is common to use more than one of these methods. For example, the job analysts may tour the job site and observe workers performing their jobs. During the tour the analyst may collect materials that directly or indirectly indicate required skills (duty statements, instructions, safety manuals, quality charts, etc.). The analyst may then meet with a group of workers or incumbents. Finally, a survey may be administered. In these cases, job analysts typically are psychologists or have been trained by, and are acting under the supervision of, a psychologist. Saville Assessment have developed the Job Profiler questionnaire, an online measure that captures the essential features of jobs in an efficient and effective manner. Please speak to your course director to discuss job analysis and job profiling requirements.

Notes:

Section 3: Assessment Choice

Considerations for Choosing Assessments

Notes:

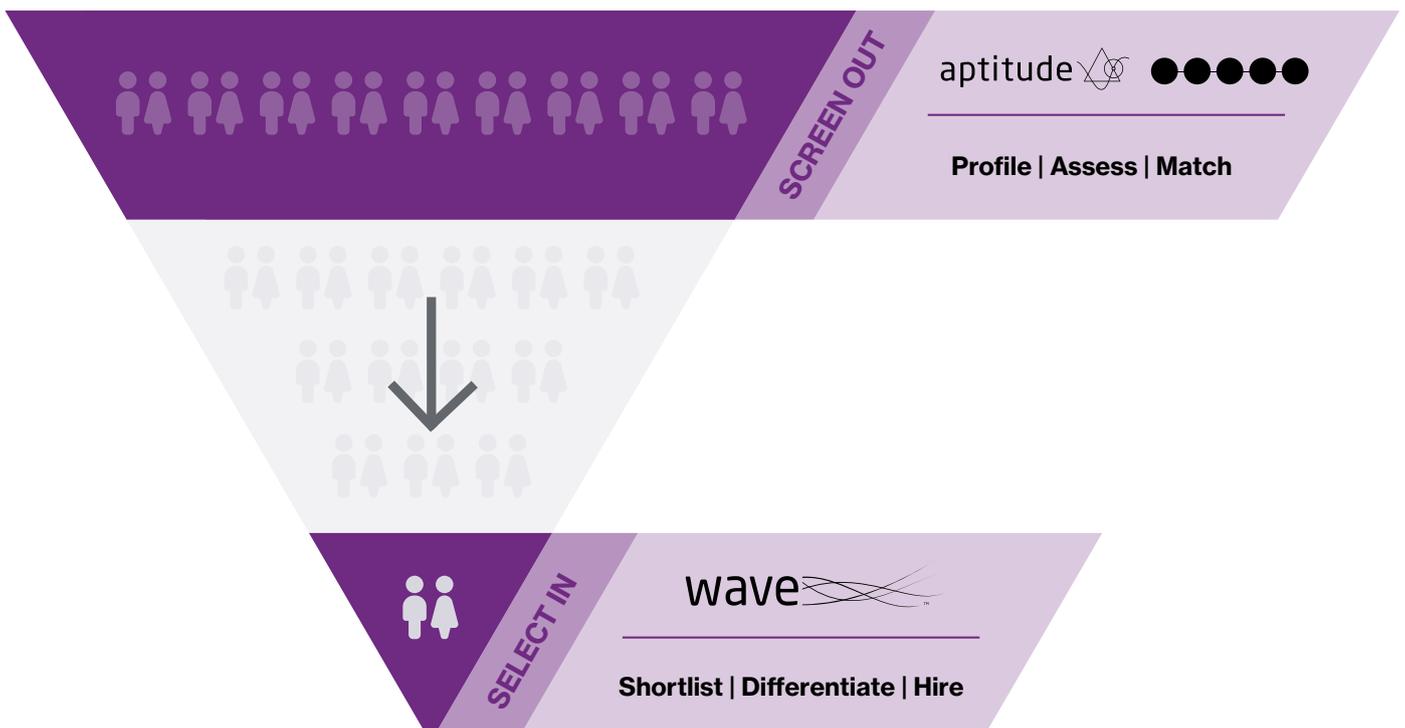
Early considerations

- Do the test yourself
 - Does it look good?
 - Does it make sense?
 - Is the content relevant to the role?
 - Does the content appear fair and inoffensive?

Other considerations (to be discussed later)

- Is it psychometrically sound?
- How much does it cost in total?
- What are the administration practicalities (screening online, supervised final stage, number of candidates, etc.)?

Screen Out, Select In



Our Methods of Screening

- Aptitude tests
- Behavioural screening questionnaires
- Language tests, e.g. Workplace English
- Situational Judgement Tests

Notes:

Saville Assessment Aptitude Tests

Analysis Range			
Target Group	Aptitudes Assessed	Norms Available	Test Options
Directors Managers Professionals Graduates Management Trainees	Verbal Numerical Diagrammatic Abstract	Senior Managers & Executives Professionals & Managers Graduates	Online Unsupervised & Supervised Swift Combination & Single

Comprehension Range			
Target Group	Aptitudes Assessed	Norms Available	Test Options
Administrative Roles Customer Service Roles Operational Roles Commercial Roles	Verbal Numerical Error Checking	Mixed Occupational Group Individual Contributors Apprentices Foundation (International Only)	Online Unsupervised & Supervised Swift Combination & Single

Technical Range			
Target Group	Aptitudes Assessed	Norms Available	Test Options
Production Roles Construction Roles Engineering Roles Scientific Roles	Spatial Mechanical Diagrammatic	Mixed Occupational Group Apprentices	Online Unsupervised & Supervised Swift Combination & Single

Swift Apprentice Aptitude			
Target Group	Aptitudes Assessed	Norms Available	Test Options
Apprentices	Verbal Numerical Error Checking Spatial Mechanical Diagrammatic	Apprentices	Online Unsupervised only

Analysis Aptitudes

Swift or single aptitudes

Verbal

1. According to Lesley Kim, which **one** of the following should a 'good leader' do?

- Allocate blame for team failures
- Create a collective sense of achievement
- Hand over power completely
- Keep pursuing projects, even if they are failing

Lesley Kim on Leadership

"To be a good leader you need to have a vision of where you are going and understand how to unite people around you in seeing that vision. I think charisma is a very important part of leadership. I have never seen a really great leader who doesn't have some kind of charisma. Great leaders know how to take risks and are comfortable taking risks. Good leaders know when to cut their losses. Good leaders pass all the glory down so that the team feels great about winning, and when things go wrong they claim responsibility. Really great leaders are willing to give the power to get things done to other people. They are not hoarders of power."

2. If Online sales volume increases by 14% next year, what will be the **approximate** volume (in million units)?

- 312
- 2,232
- 2,246
- 2,544

Annual Sales Report			
CHANNEL	Volume of Sales (Millions of Units)	Average Number of Units Purchased per Transaction	Percentage of Transactions by Repeat Customers
Retail Outlets	4,380	365	60%
Online	2,232	155	23%
Mail Order	1,512	2,100	72%
Telephone	972	270	35%

Chapter 7 Leadership

Those in authority often lack an a tend to dwell on concepts that div concepts that reflect their intercor forgetful of purpose and values th awareness of the context of their whatever it is they are responsib

1.

Panel

Operator	Effect
T	Changes shading of all figures
U	Swaps 1st and 3rd figures
V	Changes 1st figure (see illustration)

Illustration

Input	Process	Output
△ ○ ●	T	▲ ● ○
● ● ○	U	○ ● ●
○ ▲	V	▲ ▲
▲ ▲	V	● ▲

Options for question 1:

- ?
- ▲ ▲ ●
- △ ▲ ●
- ▲ ○
- ▲ ▲ ○

Diagrammatic

Abstract Reasoning Aptitude

Swift Executive or single aptitude

1.

1 2 3 4

?

●

☆

▲

□

●

☆

▲

□

☆

□

■

▲

●

Abort ■

Finish ●

Next ►

© 2016 Willis Towers Watson. All rights reserved.

Comprehension Aptitudes

Swift or single aptitudes

Verbal

< > 📖
🔄 📄 + 📄

1
2
3
4

SETTING UP YOUR ANSWER MACHINE MESSAGE FOR THE NEW TELEPHONE SYSTEM

To divert an extension to its own answer machine message:

- Go to the handset for that extension, pick up the handset, press the "Feature" button and dial "984".
- The telephone display will indicate "471" - this is the answer machine extension.

To record an answer machine greeting:

- Press the "#" key followed by "3" to record a greeting.

To pick up messages:

- The red light on the telephone handset will indicate a message.
- Press the "#" key followed by "3" to pick up a message.

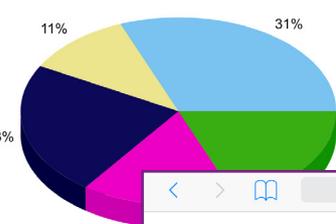
1. Which **one** of the following would best replace 'an extension' while maintaining the meaning of the passage?

- a telephone
- an addition
- an expansion
- a switchboard

< > 📖
🔄 📄 + 📄

1
2
3
4

Monthly Review of Staff Time by Activity
(Total Number of Days: 2000)



■ Sales ■ Administration

1. What percentage of time was spent on Training and Administration?

- 16%
- 19%
- 25%
- 35%
- 42%

< > 📖
🔄 📄 + 📄

1
2
3
4

Subscription Details

Company Name	Serial Number	Type*	Date
Account Manager Services	114321	G	03/05
Information Services	719909	N	09/09
Office Solutions	132461	N	12/10
Advanced Solutions	910073	P	08/11
Softwarehouse	100333	G	01/10
Marketing Focus	618736	P	06/08

*Coding Key for Type:
 P = Personal N = National G = Global

Web-Log SoftCo

Company Name	Serial Number	Type*	Date
Marketing Focus	618736	Personal	08/06

1. Please select the option, or options, which apply.

- The entire item is correct
- There is an error in the Company Name
- There is an error in the Serial Number
- There is an error in the Type
- There is an error in the Date

Abort
Finish
Next

© 2016 Willis Towers Watson. All rights reserved.

Numerical

Error Checking

Technical Aptitude

Swift or single aptitudes

Spatial

1. Which one of the objects is different in shape from the other three?

A

B

C

D

Panel description: A window with a question and four 3D rectangular objects labeled A, B, C, and D. Object A has a notch on the right side. Object B has a notch on the left side. Object C has a notch on the right side. Object D has a notch on the left side.

1. As the handle is pulled, in which direction will the end of the bar move?

Direction A

Direction B

Neither direction

Panel description: A window showing a mechanical linkage system. A handle is being pulled upwards. The linkage consists of a horizontal bar connected to a vertical bar, which is pivoted to a base. Arrows labeled A and B indicate possible directions of movement for the end of the horizontal bar.

Mechanical

1.

Panel description: A window with a logic puzzle. It features a 'Panel' with operators and their effects, an 'Illustration' table, and a question area with a sequence of shapes and a question mark.

Operator	Effect
#	Changes shading of triangles
+	Changes size of circles
@	Changes light figures (see illustration)

Input	Process	Output
○ ● ▲	#	○ ● ▲
○ ● ▲ ▲	+	○ ● ▲ ▲
● ○ ▲	@	● ○ ▲
▲ ▲ ●	@	▲ ▲ ●

Question: ? → [Operator] → [Sequence of shapes]

Options:

- ● ▲ ●
- ● ▲ ●
- ● ▲ ●
- ▲ ● ○

Diagrammatic

Swift Apprentice Aptitude

Verbal

SETTING UP YOUR ANSWER MACHINE MESSAGE FOR THE NEW TELEPHONE SYSTEM

To divert an extension to its own answer machine message:

- Go to the handset for that extension, pick up the handset, press the "Feature" button and dial "984".
- The telephone display will indicate "471" - this is the answer machine extension.

To record an answer machine greeting:

- Press the "*" key followed by "3" to record a greeting.

To pick up messages:

- The red light on the handset will indicate a message.
- Press the "*" key followed by "3" to pick up a message.

1. Which one of the following would best replace "an extension" while maintaining the meaning of the passage?

a telephone

an addition

an expansion

a switchboard

Numerical

Monthly Review of Staff Time by Activity (Total Number of Days: 2000)

1. What percentage of time was spent on Training and Administration?

16%

19%

25%

35%

42%

Error Checking

Company Name	Serial Number	Type*	Date
Account Manager Services	114321	G	03/05
Information Services	719909	N	09/09
Office Solutions	132461	N	12/10
Advanced Solutions	910073	P	08/11
Softwarehouse	100333	G	01/10
Marketing Focus	618736	P	06/08

*Coding Key for Type:
P = Personal N = National G = Global

1. Please select the option, or options, which apply.

The entire item is correct

There is an error in the Company Name

There is an error in the Serial Number

There is an error in the Type

There is an error in the Date

Spatial

1. Which one of the objects is different in shape from the other three?

A

B

C

D

Mechanical

1. As the handle is pulled, in which direction will the end of the bar move?

Direction A

Direction B

Neither direction

Diagrammatic

Operator	Effect
#	Changes shading of triangles
+	Changes size of circles
@	Changes light figures (see illustration)

1. ?

[Option 1]

[Option 2]

[Option 3]

[Option 4]

Why Use Aptitude Tests?

Hire

- Aptitude tests are mostly used for recruitment purposes, either in screening or selection

Build

- Tests are used less frequently for individual development, although career guidance and planning tools often contain an aptitude component
- They do predict training performance!

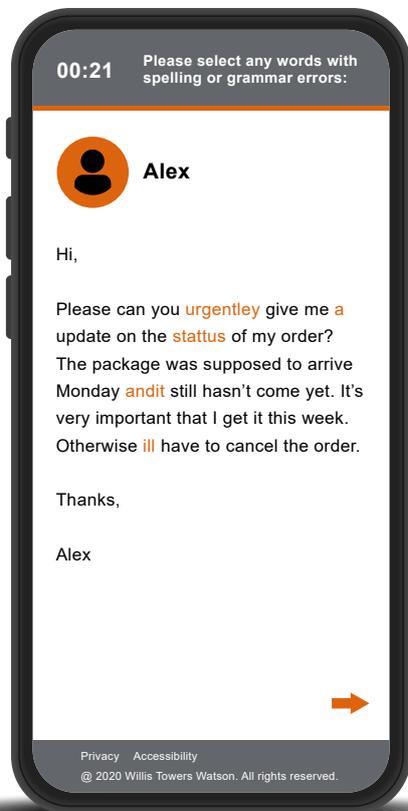
Lead

- Despite many leaders' avoidance of testing, cognitive ability has been shown to be especially predictive of performance at senior levels

Notes:

Other Screening Assessments

- Short behavioural assessments that can provide one fit score for rapid decision making in screening
- Especially effective when used together with aptitude tests
 - High validity of different criteria to complement tests
- Ideal for large-scale volume projects, merit lists can be created
- Overall fit scores calculated with data from multiple assessments
- Accurately identify the right people for the right roles based on behaviours which successfully predict performance
- Improve candidate experience with short completion times and work relevant questions showcasing strengths



- Customer Check is a new online assessment identifying how well individuals will interact with customers and represent your brand on live web chats.
- Highly-relevant assessment specifically designed to replicate typical customer contact.
- Compatible with a range of devices, including laptops, tablets and mobiles.
- Quick and engaging assessment with a 10-minute completion time.

Match 6.5 Candidate Report

MATCH 6.5

Powered by wave 



Candidate Report
Chris Park



MATCH 6.5

Powered by wave 

OVERVIEW

This report is based on your completion of the Match 6.5 questionnaire. It provides an insight into your behavioural effectiveness at work.

Using this Report

In this report you are presented with your top two talent areas. For each area, a description is provided which outlines your talent area and the special contribution that you bring to the workplace. The following information is then provided for each area:

-  **Knowing Your Talent**
Describes your talent and the benefits you bring.
-  **Making More of Your Talent**
Offers advice on how to deploy your talents for maximum impact.
-  **Using Your Talent Well**
Provides information on how over-reliance on your talents might become disadvantageous.
-  **Your Culture / Environment Fit**
Outlines which organisational cultures are likely to align with your values and needs, which in turn should lead to job satisfaction.

Your Talents

Your top two talent areas are:

Creating Innovation

Driving Success

Candidate-Friendly Feedback

Predicted Culture/Environment Fit

Based on extensive Saville Assessment research linking work place culture and the styles of individuals, this report highlights the aspects of the culture, job and environment that are likely to enhance or inhibit Sample Candidate's success:

Performance Enhancers

- ⊕ where the development of theoretical ideas and concepts is encouraged
- ⊕ where the ability to get rapidly to the core of issues and readily identify solutions to problems is highly valued
- ⊕ where there is the opportunity to take on leadership responsibilities and have control over other people and resources
- ⊕ where there is a strong results focus and determination to succeed, no matter what, and people are rewarded for achieving outstanding results

Performance Inhibitors

- ⊖ where there is little interest in the application of theoretical ideas and models and people are given little time to explore different options and possibilities
- ⊖ where little value is placed on providing new insights and identifying potential improvements
- ⊖ where there is little opportunity for taking on leadership responsibilities or directing other people
- ⊖ where the urge to achieve outstanding results is not great and people seldom persist in the face of difficulties

Option for Later Stage Interviews

Interview Scores Summary				
✓	Page	Area Assessed	Assessment Score	Interview Score
<input type="checkbox"/>	5	Evaluating Problems Examining Information (8); Documenting Facts (6); Interpreting Data (6)		
<input type="checkbox"/>	6	Investigating Issues Developing Expertise (6); Adopting Practical Approaches (3); Providing Insights (10)		
<input type="checkbox"/>	7	Creating Innovation Generating Ideas (10); Exploring Possibilities (9); Developing Strategies (9)		
<input type="checkbox"/>	8	Building Relationships Interacting with People (5); Establishing Rapport (3); Impressing People (8)		
<input type="checkbox"/>	9	Communicating Information Convincing People (8); Articulating Information (5); Challenging Ideas (10)		
<input type="checkbox"/>	10	Providing Leadership Making Decisions (10); Directing People (8); Empowering Individuals (6)		
<input type="checkbox"/>	11	Showing Resilience Conveying Self-Confidence (7); Showing Composure (7); Resolving Conflict (4)		
<input type="checkbox"/>	12	Adjusting to Change Thinking Positively (7); Embracing Change (6); Inviting Feedback (3)		
<input type="checkbox"/>	13	Giving Support Understanding People (3); Team Working (2); Valuing Individuals (3)		
<input type="checkbox"/>	14	Processing Details Meeting Timescales (2); Checking Things (3); Following Procedures (1)		
<input type="checkbox"/>	15	Structuring Tasks Managing Tasks (1); Upholding Standards (2); Producing Output (4)		
<input type="checkbox"/>	16	Driving Success Taking Action (8); Seizing Opportunities (8); Pursuing Goals (9)		

Interview Questions		
Building Relationships Interacting with People (5); Establishing Rapport (3); Impressing People (8)		Average higher potential than about 40% of the comparison group
<input type="checkbox"/> Who have you had to build a really effective, important work relationship with?		
<ul style="list-style-type: none"> • Why was it important? • What did you do to build the relationship? • How quickly did you build rapport? • How effective was the first impression you created? • How have you maintained contact? • What do you enjoy about working with new people? 		
<input type="checkbox"/> When have you had to build rapport quickly at work?		
<ul style="list-style-type: none"> • Why was it important to build rapport? • What did you do to make people feel welcome? • What did you do to put other people at ease? • What worked less well? • What lasting relationships have you developed through work? • What do you find most difficult about approaching new contacts? 		
<input type="checkbox"/> Describe a situation where you have initiated a new work contact.		
<ul style="list-style-type: none"> • Why did you choose to initiate this contact? • How did you first establish contact? • What did you do to develop the relationship further? • What were you both enthusiastic about? • What was the benefit of this relationship for your organisation? • How do you feel about networking with new people? 		

Option for Onboarding

Your Core Strengths

Communicating Information

You are prepared to put your views across confidently, persuasively and with conviction. This may be an advantage in that your opinions are more likely to be taken into account and you are likely to be able to bring other people round to your point of view.

Building your Strengths

- Before attempting to persuade others on an important issue, take time to rehearse the likely objections and best counter arguments.
- Check others' understanding of the key points you are communicating. Where you identify there is a lack of understanding, consider alternative ways of explaining the point.
- Ensure that you are being positive about the points you agree with while challenging the points you disagree with.

Working with your New Organisation/Manager

- Look for opportunities where you can make a difference through persuasion and negotiation.
- Take early opportunities to present information directly to colleagues and/or clients.
- Understand and discuss when and where it is likely to be appropriate for you to be more challenging and what the expectations are in terms of expressing concerns or alternative view-points.

Using your Strengths Well

- Could you ever find yourself continuing to persuade others when the case has already been won? **ACTION:** Observe people and listen for signs of agreement. Remember to stop persuading when these signs are clear.
- Do you often find yourself as the person who is asked to do the talking? **ACTION:** Consider when a colleague knows more about a subject and give them the opportunity to present.
- Have you ever prolonged discussion or debate by revisiting points which are agreed upon and are not going to change? **ACTION:** Know when to move on to challenge points that you have the potential to influence.

Your Challenge Areas

Structuring Tasks

You do not see yourself as particularly well organised in how you structure your work. You are more likely to take a flexible approach to managing tasks and activities rather than being particularly focused on following a precise plan or schedule.

Developing your Challenge Areas

- At the start of each day, spend time checking your diary. Understand what your upcoming commitments are and what you need to arrange to ensure that these all run smoothly.
- Reflect on recent experiences you have had which have challenged your thinking around ethics and principles. What lessons can you draw from your experiences for the future?
- Think about when you perhaps use your time less constructively and try to do something which delivers a tangible outcome instead.

Working with your New Organisation/Manager

- Prioritise and develop clear plans for work tasks and projects.
- Ensure that you understand the organisation's procedures about ethics and confidentiality, and particularly those that relate directly to your work.
- Structure your workload to ensure that it is manageable and that you are not doing several different things at the same time.

Your Actions

What actions might you consider to develop this area in your new role?

Workplace English

- Assesses candidates' ability to understand workplace English
- Secure and easily administered online in less than 20 minutes
- Tests for five different industry sectors:
 - Customer Service, Healthcare, Hospitality, Office, Operations
- Each test comprises 33 items:
 - 24 industry-specific items
 - Nine general workplace items

I [?] review the situation tomorrow

1. am
2. will
3. do
4. think

Situational Judgement Tests



Case Study: Assessment Choice

Having completed your card sort you can now start to design your assessment process. Use the steps below to help you.

1. Complete the table below, listing your seven key competencies and identify how you might measure the required characteristics

Job Title _____

Person Specification (Key Competencies from Card Sort)	Group Exercise	Role Play	Situational Judgement Test	Behavioural Assessment	Interview (Structured)	Ability Test (include test name)

Case Study: Assessment Choice Matix: Key Points

Evidence supports only certain exercises measure classic behavioural competencies

- Competency Based Interview
- Behavioural/personality assessment
- Aptitudes load onto certain criteria such as 'Solving Problems'

Aptitude tests are the single best predictor of performance

Mechanical validity where clear decision rules and cut-offs are applied consistently outperforms human judgement in screening every time!

- Generally we would advise that hiring managers should be given pre-qualified applicants without access to previous screening assessment scores

Assessment centre exercises should be used sparingly as have low validity

- Exercises should lead to one exercise score, not separate competency scores

Situational judgement tests are engaging, reflect an organisation's brand and offer strong validity, despite not measuring competencies (they measure judgement)

2. Design your end-to-end process using the funnel diagram provided

Expected volume of application and positions available are as follows:

- Call Centre Customer Support Assistant: 50 positions available, 400 applicants expected
- Senior Web Developer: 5 positions available, 150 applicants expected

Consider:

- The number of applications you expect to receive
- The number of positions you need to fill
- The level of the role
- What is being asked of the candidate at each stage
- Logistics and practical considerations

3. Identify the specific psychometrics you would use. Use the brochures, handbooks and Saville Assessment website tool to help you choose your psychometrics.

Consider:

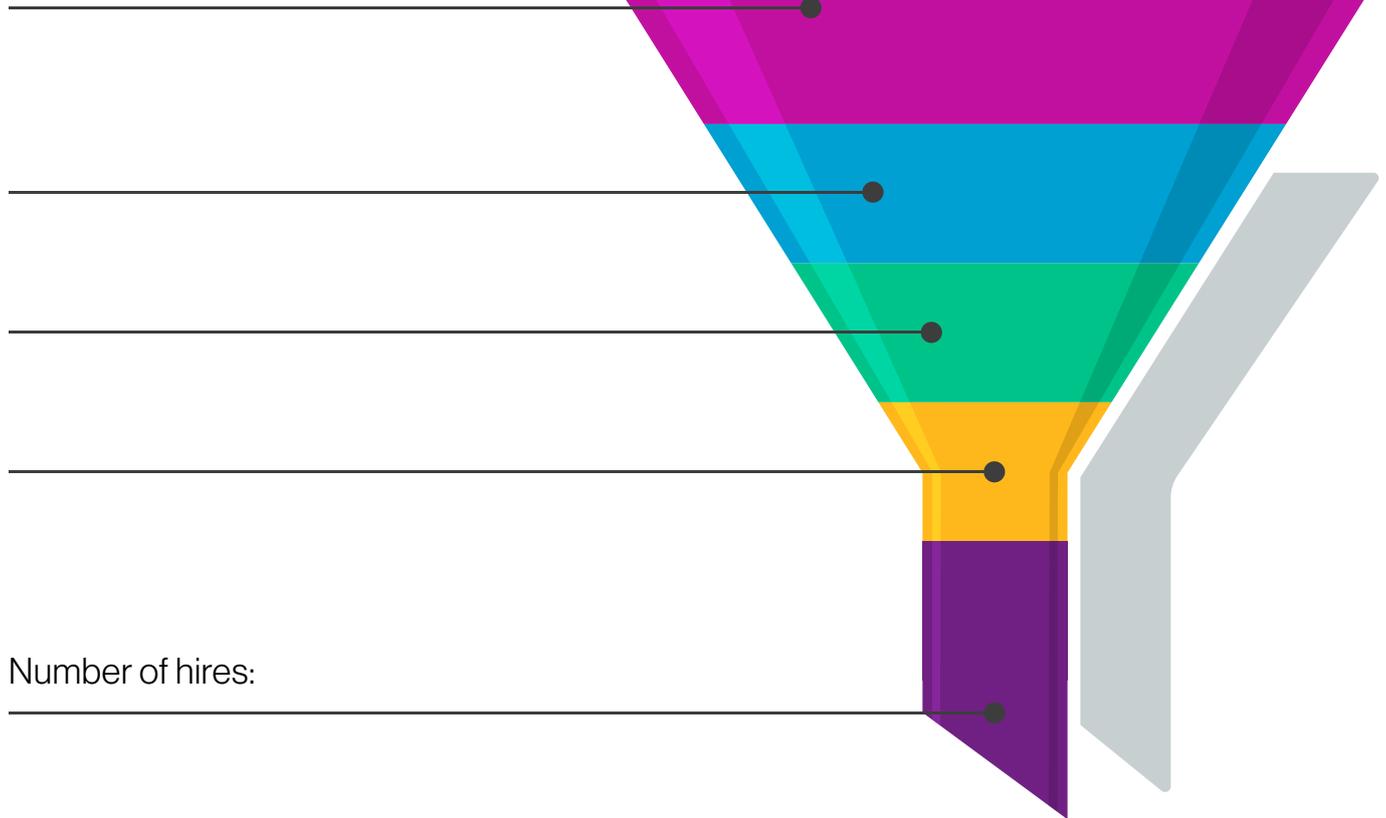
- Screening out and selecting in
- The level of the role
- The mode of administration (please specify on the funnel)
- Whether there are any reasons why a test should not be used or should be restricted*

*E.g. not appropriate for the age group, educational level, reading level, or even that the test contains content that is known to one particular cultural/ethnic group and not others.

Notes:

Assessment Process Design

Number of applicants:



Notes

Saville Assessment Aptitude Test Types

There are two types of Saville Assessment Aptitude Assessments:

Single Aptitude Tests: In-depth single tests assessing one aptitude only. Single tests consist of around 28 questions and test timings range from 6 to 24 minutes, depending on the aptitude measured. Error Checking and Spatial reasoning tend to have shorter test times, with Verbal, Numerical and Diagrammatic tests having longer test times.

As well as overall scores, users can profile scores achieved on particular types of items in the test; these are presented as 'sub-scores' on test profile charts, giving richer and more detailed information on performance.

Single tests are developed for use in both unsupervised and supervised (online) situations. For unsupervised administration (Invited Access – IA), single tests are available for each of the three ranges: Analysis, Comprehension and Technical. These tests are globally applicable across industries. Single tests for use under supervised conditions are broken down further, by key target work groups:

- The Analysis range includes Professional and Work Aptitudes
- The Comprehension range includes Operational, Commercial, Customer and Administrative Aptitudes
- The Technical range translates into Practical Aptitudes

Swift Aptitude Tests: Combination tests assessing multiple aptitudes using significantly shorter testing times than with separate single aptitude tests. A Swift test includes around 24 items with test times ranging from 9.5 to 24 minutes. Swift tests provide a mechanism to quickly establish a candidate's overall aptitude across different measures. These tests are available for use in both unsupervised and supervised (online) situations.

Swift Aptitude tests are aligned to the three ranges: Analysis, Comprehension and Technical. They are globally applicable across industries and do not break down further by work group.

Saville Assessment Aptitude Test Ranges

There are three main aptitude ranges in the Saville Assessment test portfolio: **Analysis, Comprehension and Technical**. The Analysis range has been designed for use with directors, managers, professionals, graduates and management trainees. The Comprehension range has been designed for use with operational roles in manufacturing, engineering, construction and transport, lower level commercial roles in sales, marketing, business development and financial services, customer roles in call centres, hospitality, leisure, health and education and administrative roles in private and public sector offices. The Technical range has been designed for use with production, construction, engineering and scientific roles.

Aptitude Assessment Levels

All measures within a range have been designed to have comparable levels of difficulty. They have wide bandwidth and good score discrimination, enabling deployment across a range of levels. We designated five work complexity levels with reference to the definitions of the UK National Vocational Qualifications and corresponding Zones in the US O-Net System and levels in the Stratified Systems Theory of Jacques (1996):

Level 5: Work that involves the application of a range of fundamental principles performed across a wide and often unpredictable variety of contexts, with very substantial personal autonomy and often significant responsibility for the work of others and for the allocation of substantial resources. Personal accountabilities for analysis, diagnosis, design, planning, execution and evaluation often feature.

Level 4: Work that involves the application of knowledge in a broad range of complex, technical or professional work activities performed in a variety of contexts and with a substantial degree of personal responsibility and autonomy. Responsibility for the work of others and the allocation of resources is often present.

Level 3: Work that involves the application of knowledge in a broad range of varied work activities performed in a wide variety of contexts, most of which are complex and non-routine. There is considerable responsibility and autonomy, and control or guidance of others is often required.

Level 2: Work that involves the application of knowledge in a significant range of varied work activities, performed in a

variety of contexts. Some of these activities are complex or non-routine and there is some individual responsibility or autonomy. Collaboration with others, perhaps through membership of a work group or team, is often a requirement.

Level 1: Work that involves the application of knowledge in the performance of a range of varied work activities, most of which are routine and predictable.

The graph to the right demonstrates how each aptitude range maps onto these levels.

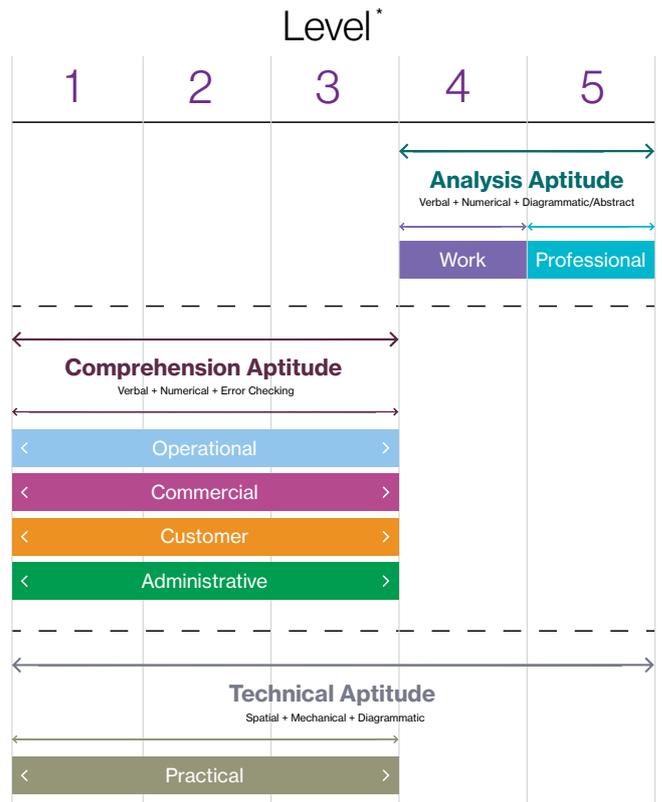
At the heart of each range sit the Saville Assessment Swift combination tests that are made up of three short sub-tests of different aptitudes. For example, the Swift Analysis Aptitude tests comprises Verbal, Numerical and Diagrammatic Analysis. There are multiple, parallel versions of each Swift assessment, with Invited Access (IA) versions for unsupervised use and a choice of online Supervised Access (SA) versions for secure verification. Each range also features a parallel series of in-depth tests of the aptitudes included in the Swift combination tests.

Analysis Aptitude Range

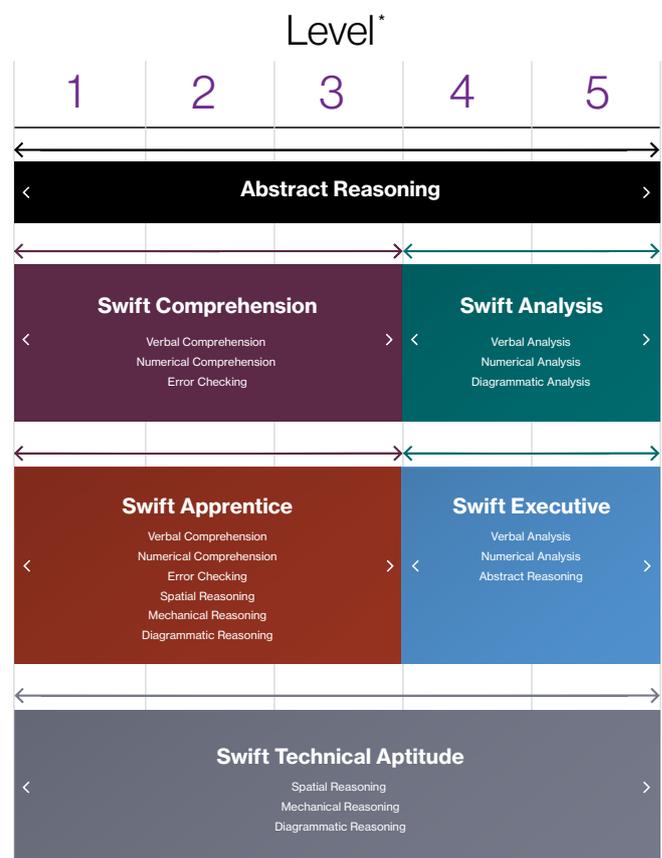
The Analysis Aptitude range comprises tests of Verbal Analysis, Numerical Analysis and Diagrammatic Analysis. It also includes Abstract Reasoning, when combined with Verbal and Numerical Analysis to create the combination test, Swift Executive Aptitude. The range's main combination test, Swift Analysis Aptitude, is available in both unsupervised (IA) and supervised (SA) formats. The range's additional combination tests, Swift Analysis Verbal and Numerical and Swift Executive Aptitude, are only available for use under unsupervised conditions (IA). Single, in-depth tests for the three key aptitudes, Verbal, Numerical and Diagrammatic, are available in both unsupervised (IA) and supervised formats (SA). For supervised formats, tests are broken down further into the key work groups Professional and Work. Professional Aptitudes has been designed for use with managers, directors and professionals. Work Aptitudes has been designed specifically for use with new graduates, trainees, technicians, team leaders and supervisors. The tests are a parallel to Professional Aptitudes but feature a wider range of content for the broader target group.

Item types for the Analysis range are as follows:

- Verbal Analysis looks at Understanding Word Meaning, Comprehending Text, Making Verbal Inferences, Evaluating Written Materials and Comparing Arguments
- Numerical Analysis looks at Understanding Tables,



* Level corresponds to NVQ (UK), O-NET (US) and Stratified Systems Theory (Jaques, 1996)



* Level corresponds to NVQ (UK), O-NET (US) and Stratified Systems Theory (Jaques, 1996)

Comprehending Graphs, Making Numerical Inferences, Evaluating Quantities and Comparing Data

- Diagrammatic Analysis looks at Understanding Logic and Rules, Comprehending Process Diagrams, Identifying Causes, Finding Faults and Comparing Flowchart Sequences

Tests available:

- Swift Analysis Aptitude – V, N and D (IA, SA)
- Swift Analysis Verbal and Numerical – V and N (IA only)
- Swift Executive Aptitude – V, N and A (IA only)
- Swift Comprehension Verbal and Numerical - V and N (IA only)
- Single Verbal, Numerical and Diagrammatic Analysis Aptitudes (IA only)
- Single Professional Verbal, Numerical and Diagrammatic Analysis Aptitudes (SA)
- Single Work Verbal, Numerical and Diagrammatic Analysis Aptitudes (SA)

Comprehension Aptitude Range

The Comprehension Aptitude range comprises tests of Verbal Comprehension, Numerical Comprehension and Error Checking. The range's combination test, Swift Comprehension Aptitude, is available in both unsupervised (IA) and supervised (SA) formats. Single, in-depth tests for the three key aptitudes, Verbal, Numerical and Error Checking, are available in both unsupervised (IA) and supervised formats (SA). For supervised formats, tests are broken down further into different key work groups: Operational, Commercial, Customer and Administrative. Each of these sub-ranges include separate tests for Verbal Comprehension, Numerical Comprehension and Error Checking, however, the content is aligned to the specific industry. Each of these tests are available online for supervised use (Supervised Access - SA).

Item types for the Analysis range are as follows:

- Verbal Comprehension looks at Understanding Word Meaning, Comprehending Text and Making Verbal Inferences
- Numerical Comprehension looks at Understanding Tables, Comprehending Graphs and Making Numerical Inferences
- Error Checking looks at Checking Letters, Checking Numbers, Checking Codes and Spotting Mistakes

Tests available:

- Swift Comprehension Aptitude – V, N and EC (IA, SA)
- Single Verbal Comprehension, Numerical Comprehension and Error Checking Aptitudes (IA only)
- Single Operational Verbal Comprehension, Numerical Comprehension and Error Checking Aptitudes (SA)
- Single Commercial Verbal Comprehension, Numerical Comprehension and Error Checking Aptitudes (SA)
- Single Customer Verbal Comprehension, Numerical Comprehension and Error Checking Aptitudes (SA)
- Single Administrative Verbal Comprehension, Numerical Comprehension and Error Checking Aptitudes (SA)

Technical Aptitude Range

The Technical range comprises tests of Mechanical, Spatial and Diagrammatic Reasoning. The range's combination test, Swift Technical Aptitude, is available in both unsupervised (IA) and supervised (SA) formats. Single, in-depth tests for the three key aptitudes, Mechanical, Spatial and Diagrammatic are available in both unsupervised (IA) and supervised formats (SA). For supervised formats, the tests are described as Practical Reasoning tests.

Item types for the Analysis range are as follows:

- Mechanical Reasoning looks at Understanding Mechanical Problems, Comprehending Physical Principles and Estimating Movement of Objects
- Spatial Reasoning looks at Estimating Lengths and Angles, Recognising Rotated Shapes and Visualising Three-dimensional Objects
- Diagrammatic Reasoning looks at Understanding Logic and Rules, Comprehending Process Diagrams and Identifying Causes

Tests available:

- Swift Technical Aptitude – M, S and D (IA, SA)
- Single Mechanical, Spatial and Diagrammatic Reasoning Aptitudes (IA only)
- Single Practical Mechanical, Spatial and Diagrammatic Reasoning (SA)

Additional Saville Assessment Tests

There are some additional assessments that either sit outside of the ranges already discussed or comprise tests across multiple ranges.

Abstract Reasoning Aptitude

Abstract Reasoning Aptitude is a 16-minute test that measures abstract reasoning. The test is available for use in both unsupervised (IA) and supervised (SA) conditions.

Swift Apprentice Aptitude

Swift Apprentice Aptitude is designed for use with operational, technical, engineering, manufacturing and construction apprentices. It combines aptitudes from the Technical and Comprehension ranges to create a combination test assessing Verbal Comprehension, Numerical Comprehension, Error Checking, Spatial Reasoning, Mechanical Reasoning and Diagrammatic Reasoning. Swift Apprentice Aptitude takes just 19.5 minutes to complete. It is available for use in unsupervised (IA) conditions only.

Saville Assessment Work Strengths

Saville Assessment Work Strengths is a short but highly valid assessment designed for use in volume recruitment across a wide range of different job roles. Work Strengths provides a highly positive approach to the selection of staff, with feedback focusing on candidates' strengths.

Work Strengths: suitable for use with graduates, management trainees, managers and professionals

The tool is based upon research integrating personality, competency and overall effectiveness at work and is aligned to both the Big Five personality facet model and the Great Eight competencies. Its development benefited from a performance-driven approach, whereby the validity of the questionnaire is maximised by selecting items from the pool that are most predictive of performance at work.

The assessment uses a dynamic response format that utilises both ratings and rankings, allowing for identification and control of distortion.

The Work Strengths output is simple to use, enabling managers across organisational levels to interpret profiles easily and accurately. It provides feedback on work culture and the environments in which a candidate is likely to be most and least suited, as well as optional interview questions.

Match 6.5

Match 6.5 is a fast and valid tool for screening high volumes of candidates; taking just 6.5 minutes to complete. It allows the client to screen with one behavioral fit score alone or a fit score combining behaviour and aptitude.

Match 6.5 easily screens across different roles for large organisations hiring across different job areas.

Saville Assessment Workplace English Tests

Workplace English tests assess an individual's ability to understand workplace-relevant sentences in English. All tests are available online through Invited Access administration. Separate online versions are also available for supervised testing.

Saville Assessment Situational Judgement Tests

Situational Judgement Tests or SJTs provide engaging, realistic, work-related previews of the role by presenting candidates with scenarios they are likely to come across on the job. Candidates are then presented with a series of response options for each scenario and asked to rate the effectiveness of each. SJTs create opportunities for impactful employer branding and offer fast and effective screening for high volumes of candidates.

Section 4: Test Administration

Administration Formats

- **Open Access** – This 'open mode' allows any individual to access the test from a publicly-accessible location. For example, candidates can be offered online access to practice tests or career guidance assessments.
- **Invited Access** – This 'controlled mode' provides secure access to specific individuals to complete online tests in an unsupervised setting e.g. a password and link are emailed.
- **Security advice for invited access administration:**
 - Ensure candidates are pre-qualified and cannot sit a test multiple times
 - Use the advice on test resets documented in your organisation's testing policy
 - Use item-banked tests, e.g. gradient step testing
 - Review candidate response patterns ('forensic analysis'), e.g. unusually fast response times
 - Retest under supervised conditions
- **Supervised Access*** – This 'managed mode' is where a trained test administrator verifies identity and initiates the test administration online. The test administrator oversees the test environment/process throughout, e.g. at a secure testing centre.

*There is another 'supervised mode' which Saville Assessment do not advocate, where the identity, login and final completion is verified by a supervisor, but the actual test completion is not managed.

Notes:

Preparing for Administration

- Review your organisation's testing policy to ensure that your communication and administration are compliant with it
- Prepare your communication (email/letter)
 - Check for any reasonable adjustment requirements and any anticipated problems completing the tests
 - Ensure candidates have access to preparation/practice materials
 - Inform candidates of how you will store their data and for how long, and who will have access to it, ensuring you are compliant with any applicable legislation
 - Inform candidate of next steps, e.g. when they will receive feedback

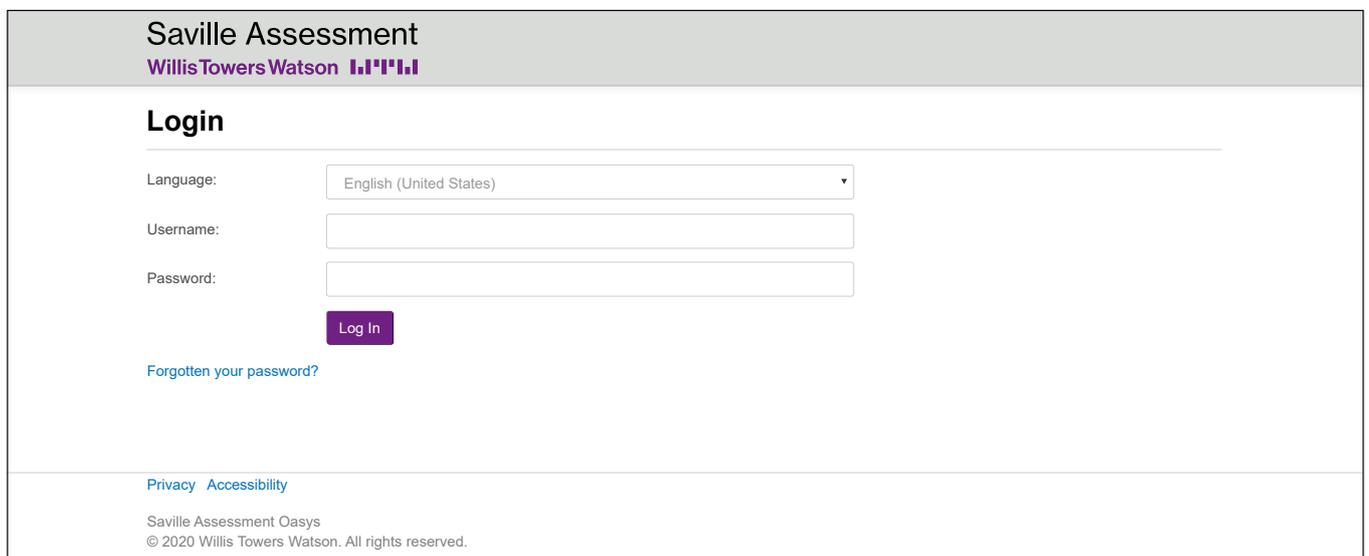
Administration via the Saville Bureau

- Complete a form – our Bureau team do the rest!
- Cost-effective for small numbers
- Fast turnaround within two hours
- Open 8am – 6pm UK time

Notes:

Administration Using Oasys

- One-off investment
- Free training to use Oasys
- Better value for larger numbers
- You are in control of your own projects
- Can be branded or non-branded
- 99.9% uptime



Saville Assessment
Willis Towers Watson

Login

Language: English (United States) ▼

Username:

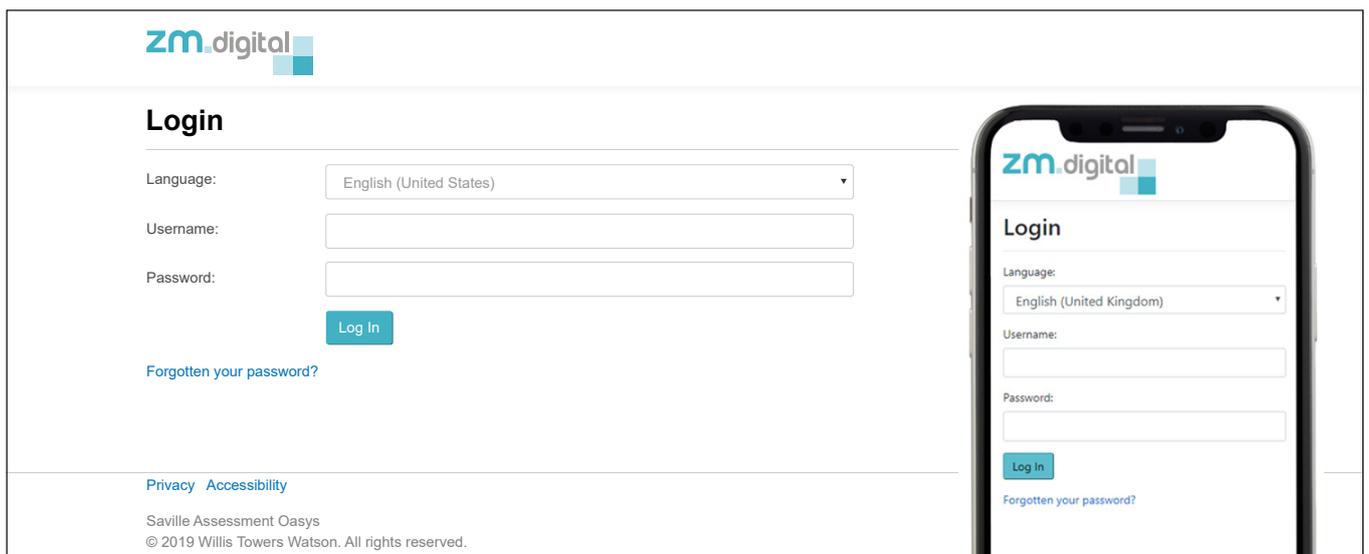
Password:

[Log In](#)

[Forgotten your password?](#)

[Privacy](#) [Accessibility](#)

Saville Assessment Oasys
© 2020 Willis Towers Watson. All rights reserved.



zm.digital

Login

Language: English (United States) ▼

Username:

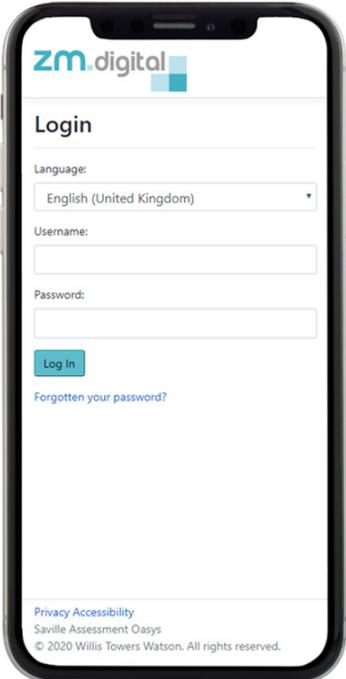
Password:

[Log In](#)

[Forgotten your password?](#)

[Privacy](#) [Accessibility](#)

Saville Assessment Oasys
© 2019 Willis Towers Watson. All rights reserved.



zm.digital

Login

Language: English (United Kingdom) ▼

Username:

Password:

[Log In](#)

[Forgotten your password?](#)

[Privacy](#) [Accessibility](#)

Saville Assessment Oasys
© 2020 Willis Towers Watson. All rights reserved.

Administering Tests via an ATS Integration



Administering Hard-Copy Assessments

- The History of Testing

Preparing the Session:

- Select appropriate tests
- Timetable of the day
- Inform the candidate: invitation letter, check accommodations, preparation guides
- Materials – order in, check re-usables
 - Question booklets, answer sheets, pencils, rough paper, calculators, administration instructions, stopwatch, test log
- Familiarity with test – materials, admin instructions, example questions
- Check testing environment

Notes:

Introducing the Session:

- Introduce yourself
- Rationale for testing
- Description of tests used
- Length of test session
- Feedback arrangements
- Confidentiality, including data storage (under applicable legislation) and who has access, i.e. authorised test users
- Informed consent
- House keeping points: reading glasses, phones, toilets
- Testing conditions
- Refer to admin card and example questions
- Any questions

Administering the Session:

Follow administration instructions precisely!

- Distribute candidate materials but not test booklets
- Read clearly and slowly
- Ensure personal details are filled in
- Help individual candidates to get examples correct
- Use a stopwatch and independent time
- Keep your eyes on the room throughout test
- Ensure everyone has stopped
- Collect all materials
- Thank the candidates

Administration Guidelines

Below are some guidelines for use when you are running an administration session. Whether you're administering a hard-copy or online testing session, a lot of the preparation and introduction will be similar.

Before the Session:

Provide the participant with as much information as you can to ensure that they have given their informed consent. This should include most of the points given below – especially if you are inviting them to complete an online assessment where they will not have a formal administration. See also example invitation letter.

Ensure that you have any relevant background information and make any necessary alterations to the room/layout/schedule. Check the suitability of the room – lighting, space, seating, layout, disruptions.

To Introduce the Session:

- Welcome the candidates
- Introduce yourself – name, job role
- Be professional and friendly – try to put candidates at ease
- Describe the assessment process, (e.g. tests, interview, site tour)
- Explain how many tests candidates will be completing and include breaks (if more than one test)
- Explain why the test(s) is/are being used
- Explain how the data will be used and stored (in line with applicable legislation)
- State how long each test will last
- State how long the whole session will last (total test time + 15 minutes per test)
- Remind them to fetch reading glasses, go to the bathroom, switch off mobile phones, test conditions apply, no conferring
- Tell them when/how they will receive feedback
- Set an appropriate pace – should work quickly but accurately
- Mention example questions – not timed or scored
- Explain you will be reading from Administration Card – ensure standardisation
- Give them the opportunity to ask questions

For the Assessment:

- Place relevant materials on desk (all except question books can be laid out beforehand)
- Hand out question booklets as instructed
- Read instructions precisely as they are written – be clear
- Explain examples appropriately, coach and support where required

Ending the Session:

- Collect all materials (count in)
- Check names are included
- Thank candidates
- Ensure candidates understand next steps
- Check materials for completion (and for marks on reusable booklets)

Notes:

Sample Introduction to Aptitude Assessment Session

Hello, welcome to....., I'm....., I am a..... with..... I will be conducting this testing session with you.

You should have all received the details describing what will happen during the testing session. I'll be asking you to complete.....test(s). You will then be having an interview/completing an in-tray exercise etc. We hope to complete the assessments by.....o'clock today.

Before you start the test(s), I'll explain why we're using them; we use them because they give us a fair and objective assessment of your skills in.....
.....**(Insert short description of relevant ability)**. These are important skills for the role you have applied for and we find that those applicants who do well in the tests subsequently do well in the role. It is also in your own interests as well as ours that you are suited to the role for which you've applied. We also get additional information from the tests that we can't readily get from other aspects of our selection process.

Decisions on whether to progress your application are based on all the information we gather from you today; you can contact me for feedback on your test performance, and I'll give you my contact details later.

Your results from this session are confidential and will be stored in line with applicable legislation. Your results will only be shared with those involved in the recruitment process; do I have your informed consent to continue?

The whole test session will last approximately.....minutes **(add on 15 minutes to each test duration)**. The first test will last for exactly..... minutes. Please don't leave the room once we're underway with the tests, so now is a good time to visit the toilet or collect reading glasses if you need them; please do not talk to other candidates once we've started the tests and please can you ensure mobile phones are switched off.

I'd like to take this opportunity to advise you to work quickly through the tests and try to answer as many questions as possible. I'll be reading the instructions for the test to you from a card, so please listen carefully. There are some example questions at the beginning of each test so you know what to do on the test itself. These are not timed or scored and I'll walk round to ensure that everyone has got them correct before we continue – they don't contribute to your overall score.

Are there any questions before we get started?

Notes

Saville Assessment Test Formats

Invited Access (IA): This online format is available to Oasys and Bureau users, and presents one of many parallel versions assembled from content that is completely separate from the supervised versions. The self-administering, interactive format in conjunction with economical time limit makes this tool ideal for unsupervised assessments. It can be taken in any place with a computer and good internet connectivity (e.g. home, office, library, internet café). The assessment tool can be deployed as a standalone tool in talent audit and self-development situations, or as a screening tool in a multistage recruitment process. For 'high-stakes' appointment decision-making, Saville Assessment strongly recommend verification of unsupervised test results through supervised assessments.

Supervised Access (SA): Due to its self-administering, interactive format it is ideal for use when regularly testing individuals in a specific location with computer access and good internet connectivity, and with groups of candidates in dedicated testing centres. The assessment is particularly useful in advance of line manager or panel interviews, and can be used as a stand-alone tool or to crosscheck results of an unsupervised online assessment.

Saville Assessment's Bureau Service

Candidates can complete Saville Assessment Wave and Aptitude assessments in an unsupervised, 'Invited Access' environment. A secure link to the assessment platform is emailed to the individual along with a password and username. To invite a candidate to complete a set of tools, administrators contact Saville Assessment's Bureau Service in advance of the administration. Saville Assessment's Bureau team require the first name, last name and email address of the candidate and the trained user. Following completion of the assessment, the required reports are generated automatically and can be viewed by the appropriate party. Aptitude assessment reports have been designed for test users, candidates and line managers.

Administrator Tasks

Inform the candidate by letter or email of what is involved, i.e. the types of exercises they will be given, the length of time each exercise will take and anything they may need to take with them. A reminder to bring reading glasses is often advisable if taking aptitude tests and other written assessments.

Ask the candidates whether they have any need for accommodation as a result of disability e.g. vision, hearing, physical, etc. If a candidate notifies you they are dyslexic then you may decide to have a policy that asks the candidate to give certified evidence of their dyslexia. If the candidates do inform you of their disability, you may well need to follow up to establish what an appropriate accommodation would be. See discussion in the Best Practice & Ethics chapter.

Best practice is to **send all candidates preparation guides**, which can be downloaded from the Saville Assessment website free of charge or can be purchased on paper. This creates a more level playing field between candidates who have in the past had more and less exposure to tests.

Complete a Bureau Request Form and send to Saville Assessment's Bureau Service.

Saville Assessment's Oasys Platform

As an alternative to using Saville Assessment's Bureau, trained users can purchase their own Oasys platform, which allows for Wave and online ability test administration and reporting, independent of Saville Assessment's Bureau Services. Oasys platforms can be tailored to include client branding and logos, and can be incorporated with clients' existing online platforms, offering an integrated online assessment process.

Administrator Tasks

Inform the candidate by letter or email of what is involved, i.e. what are the types of exercises they will be given, the length of time each exercise will take and anything they may need to take with them. A reminder to bring reading glasses is often advisable if taking aptitude tests and other written assessments.

Ask the candidates whether they have any need for accommodation as a result of disability e.g. vision, hearing, physical, etc. If candidate notifies you they are dyslexic then you may decide to have a policy that asks the candidate to give certified evidence of their dyslexia. If the candidates do inform you of their disability, you may well need to follow up to establish what an appropriate accommodation would be. See discussion in Best Practice & Ethics.

Best practice is to **send all candidates preparation guides**, which can be downloaded from the Saville Assessment website free of charge or can be purchased on paper. This creates a more level playing field between candidates who have in the past had more and less exposure to tests.

Set up an Oasys project and **add candidates to this project on Oasys**.

Online Supervised Administration

Any PCs, laptops or tablets that will be used to administer online assessments must have internet and email access. Administration is secure in that a trained test administrator must be present to administer the assessment and to allow the candidate access to the tools via the internet. The online assessments have self-explanatory instructions and examples to ensure all candidates receive the same information, so there is no need for an administrator to read these out.

Administrator Tasks

Inform the candidate by letter or email of what is involved, i.e. what are the types of exercises they will be given, the length of time each exercise will take and anything they may need to take with them. A reminder to bring reading glasses is often advisable if taking aptitude tests and other written assessments.

Ask the candidates whether they have any need for accommodation as a result of disability e.g. vision, hearing, physical, etc. If candidate notifies you they are dyslexic then you may decide to have a policy that asks the candidate to give certified evidence of their dyslexia. If the candidates do inform you of their disability, you may well need to follow up to establish what an appropriate accommodation would be. See discussion in Best Practice & Ethics.

Best practice is to **send all candidates preparation guides**, which can be downloaded from the Saville Assessment website free of charge or can be purchased on paper. This creates a more level playing field between candidates who have in the past had more and less exposure to tests.

Check the room/environment ideally around half an hour before you start testing. Is the room a comfortable temperature? Check if there is a lot of noise coming from any source. Is there water for you and candidates? Does the technology work?

Put a sign on the room door to inform people to be silent and not to disturb as testing is in progress. If you are in a busy corridor where people congregate to talk, you may

want to put sign on a flipchart in the corridor.

Turn your own mobile phone/tablet off!

Introduce the session, providing detail on what tests the candidates will be doing and why and how long the session will last.

Ensure you **keep your eyes on the room** throughout the test. You may choose to walk around checking that all is in order every 10 minutes or so.

Supervised Administration

Administrator Tasks

This section relates to the history of testing, using hard copy assessments. However, many of these principles remain relevant across all versions of supervised assessment.

Inform the candidate by letter or email of what is involved, i.e. what are the types of exercises they will be given, the length of time each exercise will take and anything they may need to take with them. A reminder to bring reading glasses is often advisable if taking aptitude tests and other written assessments.

Ask the candidates whether they have any need for accommodation as a result of disability e.g. vision, hearing, physical, etc. If candidate notifies you they are dyslexic then you may decide to have a policy that asks the candidate to give certified evidence of their dyslexia. If the candidates do inform you of their disability, you may well need to follow up to establish what an appropriate accommodation would be. See discussion in Best Practice & Ethics.

Best practice is to **send all candidates preparation guides**, which can be downloaded from the Saville Assessment website free of charge or can be purchased on paper. This creates a more level playing field between candidates who have in the past had more and less exposure to tests.

You should also **become familiar with all the testing materials**, paying particular attention to reading the administration card and really getting a deep understanding of the **example questions**.

Prepare a list of all the materials you need several days before the testing session and make sure you have enough materials. Having enough booklets and answer sheets is critical. For each candidate you will need:

- Booklets + two spare (check they have not been written on!)
- Answer sheets + two spare

Section 4: Test Administration

- Pencils (at least two per candidate and spares)
- Erasers – if required
- Calculators – if required (and spares – the batteries all seem to go at the same time!)
- Scoring keys – if required
- Scrap paper – remember to collect at end of session

And at least one for each test (for administrator use):

- Pencil sharpener
- Stopwatch
- Separate watch with second hand to allow you to take an independent check of the time

Check the room/environment ideally around half an hour before you start testing. Is the room a comfortable temperature? Check if there is a lot of noise coming from any source. Is there water for you and candidates?

Arrange the tables and chairs. A large desk at the front of the room to hold all the materials for the administrator is always useful. Various arrangements of desks can work depending on the number of candidates. You are trying to ensure that the arrangement enables you, the administrator, to make eye contact with every candidate, that candidates are well spaced and not looking at another's answer sheet. A herringbone or V-shaped pattern often is preferable to a standard 'rows and columns' classroom format.

Has each person got enough space for themselves? It is often a good idea to lay out materials in one of the positions the candidate will be sitting at to check if there is enough space for all the materials. Is the room bright enough? Is there a glare from any light source which makes the booklets difficult to read?

Provide a wall clock in a position that all the candidates can see.

Put a sign on the room door informing people to be silent and not to disturb as testing is in progress. If you are in a busy corridor where people congregate to talk, you may want to put sign on a flipchart in the corridor.

Turn your own mobile phone/tablet off!

Introduce the session, providing detail on what tests the candidates will be doing and why and how long the session will last.

Follow administration instructions exactly. It is tempting to omit or add words but this should be avoided to ensure

consistency and clarity.

Read clearly and slowly, and at a good volume from the card.

Ensure personal details are completed on the answer sheet for every candidate (the answer sheet will not be of much use if you do not know who it belongs to!)

Check if **example questions are correct**. Point out individually to each candidate if they have got one wrong. Then ask them questions to help work out how they are in error and help them to see why the answer they have chosen is incorrect and what the correct answer is.

Use a stopwatch, but also **take a note of time independently**, which if your stopwatch fails will give you a backup.

Ensure you **keep your eyes on the room** throughout the test. You may choose to walk around checking that all is in order every 10 minutes or so.

At the end of the test make sure that everyone has stopped and then **collect all the materials** ensuring you have all the answer sheets, booklets and rough paper.

Test User and Test Administrator Roles

Test User: Occupational Ability trained individuals need to carry out job analysis or competency requirement profiling and as a result choose appropriate tests to utilise.

Test Administrators trained to the BPS certification standard can then lend a hand by administering, scoring and norming the tests.

Interpretation of results in the light of reliability, validity and job requirement information as well as decision-making are restricted to Test User: Occupational Ability trained test users.

Notes:

Section 5: Test Scores

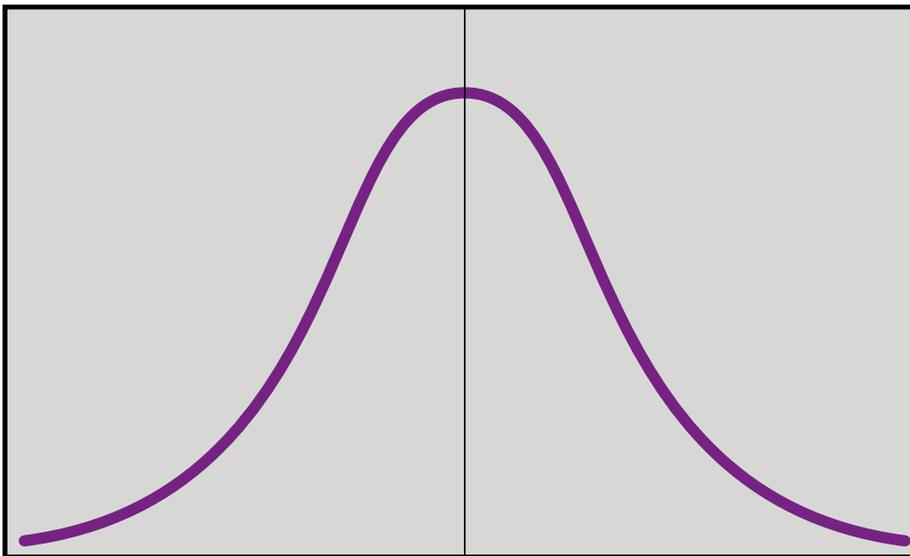
Scoring

Notes:

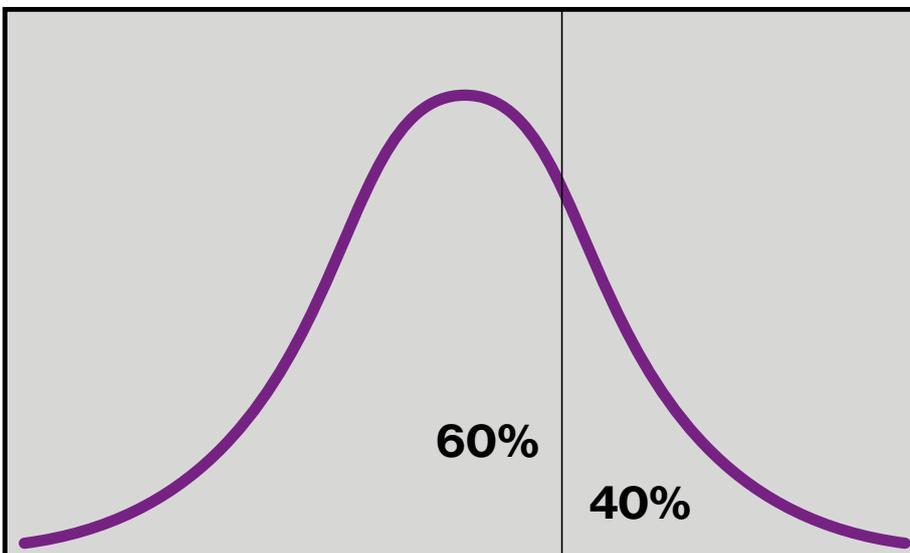
An applicant scores 19 correct (their raw score) on a numerical test (28 items)

How well has the applicant performed?

The Normal Distribution



Percentiles

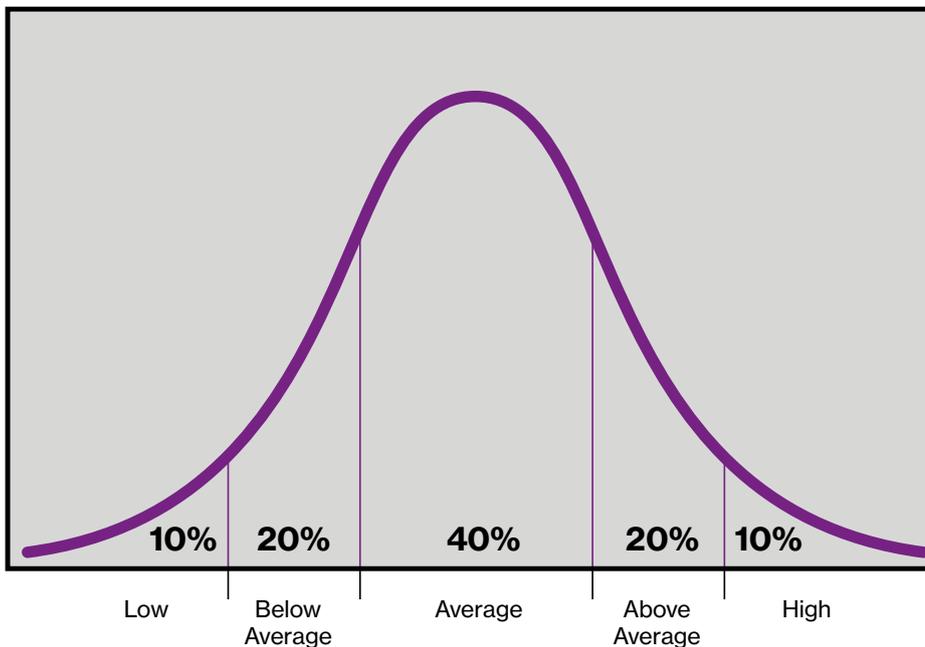


Describing Percentiles

Notes:

- Candidate A: 56th %ile Average
- Candidate B: 33rd %ile Average
- Candidate C: 68th %ile Average

The Normal Distribution Performance Bandings



Scoring by Hand

- Check the answer sheet for any odd patterns
- Follow the instructions for scoring carefully
- Count the number attempted, the number correct, sub and total scores
- Double check the scoring
- Convert raw scores to norm scores – Percentiles, Stens, T-scores

Notes

In this section of the course notes, we will introduce you to test scores and answer the question “Why use test norms?” We will be looking at tests norms in more detail later on in the course. For now, we will become familiar with the normal distribution, a key concept in test norms and test score interpretation and understanding.

Raw Scores Vs Normed Scores

For a score to be meaningful to us, we need some way of comparing the score achieved by an individual on a test against the scores achieved by a representative/relevant sample of people, i.e. a benchmark or norm group. To start with, consider the following example:

An applicant scores 19 correct answers (their raw score) on a Saville Assessment Professional Numerical Analysis test, which has 28 questions (also known as items) and lasts for 20 minutes. How well have they performed?

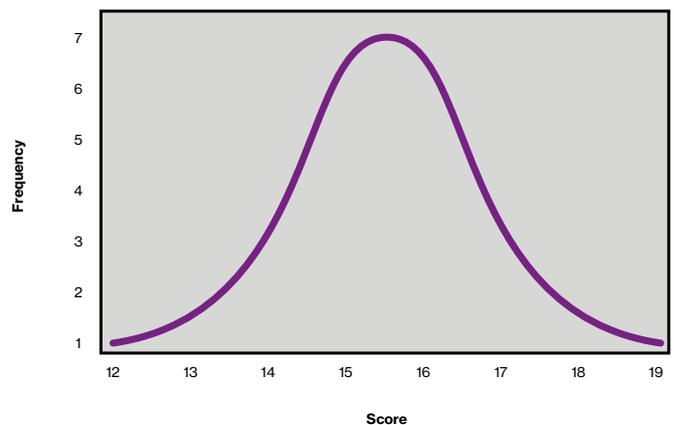
The answer is that without knowing how well the rest of the group has performed, it is impossible to know whether 19 is a good score or a bad score.

Frequency Distributions

One way to look at a group's scores is to produce a frequency distribution. On the horizontal (x) axis, the scores on the test or assessment are presented and on the vertical (y) axis, the frequency (or count) is presented. The frequency count for each score is plotted on the graph to give us our frequency distribution.

The Normal Distribution

Scores which, when plotted, form a smooth curve like the one depicted below are said to be ‘normally distributed’. This curve is sometimes called a ‘bell-shaped curve’. You can see that most scores fall around the average (bulge in the middle), with fewer occurrences towards the far left (low scores) and towards the right (high scores).



Most natural phenomena are normally distributed. If you were to plot the shoe sizes of a large number of women in the UK, you would discover a normal distribution. You would find the same with height and weight. Using these normal distributions, you are able to get a sense of where you stand compared to others. You can start to answer questions such as: Are my feet big, small or average? Am I tall, short or around normal? How well have I done in a numerical test? The normal distribution and its unique properties are the basis for all test norm systems.

Norm Systems

Grades/Bands

Grades or bands are a type of rank order norm system. The area under the curve represents the total percentage of people who have taken a particular test. We are able to chop up the curve into bands of average, above average, below average, high and low. Or indeed grades of A, B, C, D and E. Grades and bands are one of the simplest norm systems we can use.

Benefits of using grades or bands:

Grades and bands are easy to understand and to interpret

and hence they are often used in education to describe the performance of students.

Drawbacks of using grades or bands:

Grades and bands, by their nature, are not precise measurements. There are 'grey' areas at the borders and a single mark may mean the difference between an average and above average score. To overcome this lack of precision and differentiation, another norm system can be used: Percentiles.

Percentiles

Percentiles are essentially an extension of the Grade System – they are 'graded grades' whereby instead of having just five bands, you have many bands, giving you a more sophisticated grading system. In fact, the percentile system splits the normal distribution into 100 bands, each representing 1% of the comparison group or population under the curve. Percentile is defined as 'per cent', or 'of a hundred'.

The normal distribution can, therefore, be thought of as being divided up into percentiles. A percentile rank indicates the percentage of the norm group a person's score comes above. A score at the 60th percentile means that the individual performed better than 60% of the group (while 40% of the group have performed better than them). The way to describe a score at the 60th percentile is to say "you have performed better than 60 percent of the comparison group". This phrasing is useful when feeding back test scores to candidates or line managers.

Benefits of using percentiles:

Describing scores in terms of percentiles can be useful as they give a more precise measurement of a candidate's scores than grades or bands, enabling greater differentiation between candidates. Percentiles are commonly used with tests because they are easy to understand and to interpret.

Drawbacks of using percentiles:

However, percentiles are not equal units of measurement. An increase from the 95th to the 99th percentile is a greater performance improvement than an increase from the 50th to the 70th percentile. Percentile scores can therefore be said to reduce the differences at the extremes and exaggerate scores around the middle of the distribution. When using percentiles, it is, therefore, key that you do not over-read small differences between applicants; one or two raw scores differences near the average can result in a large gap in percentile terms.

This leads to a major practical problem - you cannot take percentiles from different tests to produce overall composite scores. In order to get around this problem, we use scores called standard scores, should an occasion arise when you wish to add or take an average of a set of scores. We will look at standard scores in detail later in the course.

Scoring by Hand

Previously, all assessment was conducted with paper and pencil assessments. With the growth of online assessment, hard copy tests and scoring by hand are now largely obsolete. However, it can be helpful to understand hand scoring to appreciate the principles behind our online scoring. This could be also be useful information to know for the post-course work if you decide to register with the BPS. Saville Assessment hard-copy test answer forms have integrated scoring keys and profile charts. The first thing to do is check for odd patterns of responses; this may include candidates missing out several pages of the test, or giving more than one answer to a question requiring only a single response.

The instructions for scoring are contained inside the answer form; simply open to access them. You'll need a calculator to work out the test taking style and sub-scores, and it is worth double checking your calculations or asking a colleague to check your workings out. Scoring is easy to do, but it is also easy to make mistakes.

Once you've calculated the total score, test taking style scores and also the sub-scores, you can profile them on the profile chart. Follow the instructions in terms of the number of T score dots that need to be drawn onto the profile chart – these can be found at the bottom of the profile chart. You can then read off the Percentile, Sten or T scores for the candidates' scores, and write them on the summary score sheet. We will learn about Sten and T scores later in the course.

As part of the feedback process, the profile chart can be detached from the answer form and discussed with the candidate. They can also take this away for future reference. There are tips for the candidate on how to improve the relevant ability on the back of the profile chart which candidates may find useful.

Section 6: Test Feedback

Purpose of Test Feedback

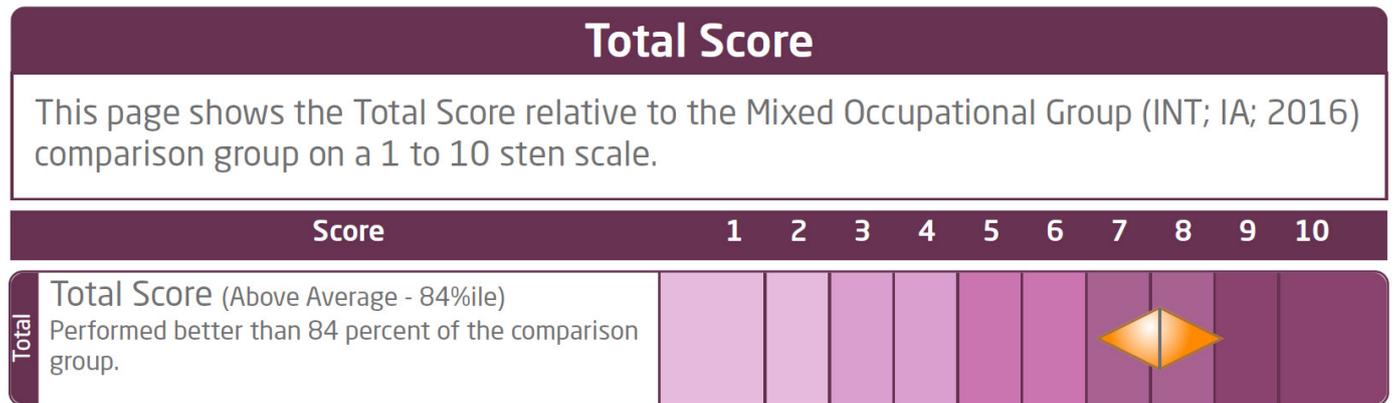
- For public relations – candidates are often clients too
- To assist in explaining why job offer not made
- To comply with applicable legislation
- To understand results by seeking examples/explanation
- To understand conflicts with other assessment data
- To increase recruited candidates' self awareness
- To develop individuals in key ability areas

Feedback of Saville Aptitude Tests

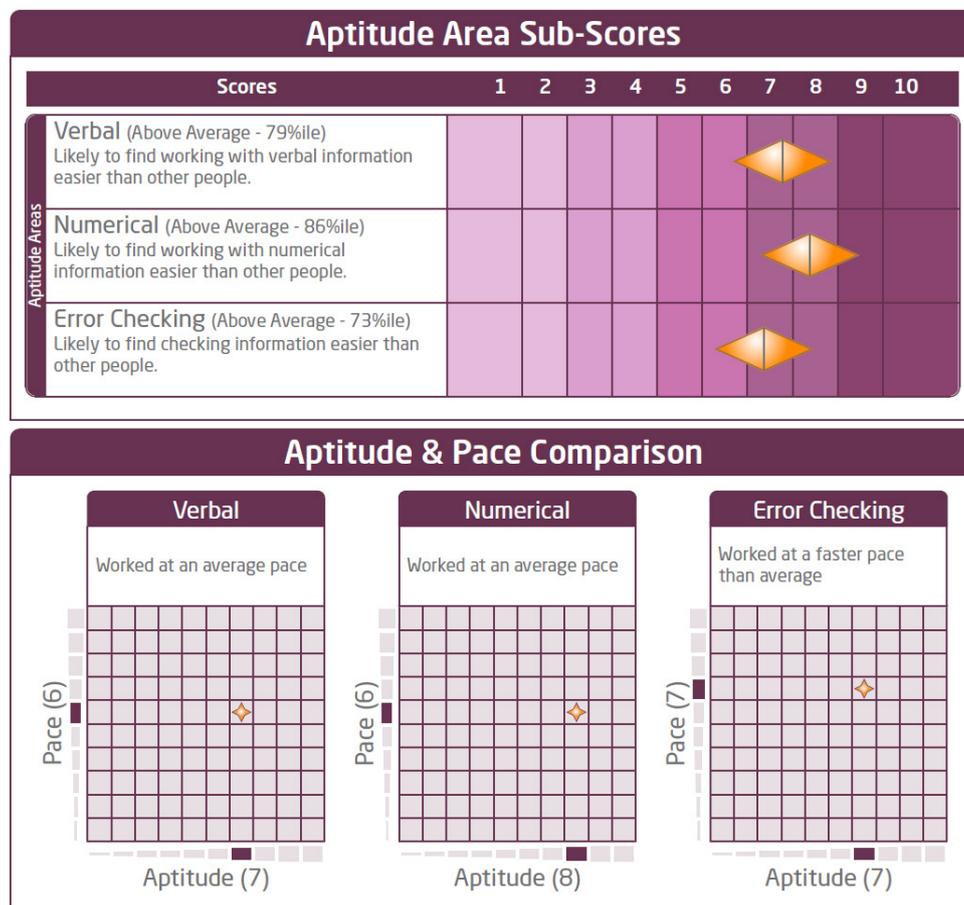
- **Total score** is main focus in feedback and decision making
- **Sub-scores** help to describe what the test is assessing
- **Item type** or **sub-scores** also help pinpoint very specific strengths and development needs
- Unsupervised online item-banked tests feature information about the combination of an individual's **Pace** and **Aptitude**
- Supervised fixed-content tests contain information about an individual's **Speed**, **Accuracy** and **Caution**
- **Feedback reports** give straightforward feedback and development advice for candidates, line managers and trained test users

Notes:

Unsupervised Item-Banked Profile Example



Unsupervised Item-Banked Profile Example



Feedback Process

- Introduction and purpose
- Summary of tests used and comparison group
- Background
- Encourage self assessment
- Discuss overall performance
- Review pace and sub-scores (for in-depth feedback)
- Summary and review

Feedback Tips – Do's and Don'ts

- Do build rapport
- Do ensure two-way dialogue
- Do gauge candidate reactions and impressions
- Do discuss development areas if raised by candidate
- Don't give too much information or too many scores at once
- Don't use technical jargon
- Don't make value judgements

Notes:

Notes

Purpose of Test Feedback

Even if candidates are not successful, research shows that they are more positive about an organisation if they have received feedback on their performance. Feedback also gives candidates a chance to understand why tests are used by the organisation and the rationale behind using them. Candidates and feedback providers can discuss the pattern of results displayed, and discussion of examples from their working life can aid understanding. Where there are differences between ability test results and findings gained from other assessment centre exercises involving numerical calculations, for example an in-tray forecasting exercise, these can be explored. Candidates can gain valuable insight into their relative strengths and development areas, which can help guide personal development in specific areas.

Candidates have the right to see any information held on them, including assessment results. It is best practice to provide this in an appropriate and accessible form, such as verbal or written feedback. Candidate feedback reports are designed for this purpose.

Feedback of Saville Assessment Tests

Saville Assessment ability tests give additional, unique information regarding test performance on different categories of questions on a test. These measures break down the overall total score into **Item Type** or **Aptitude Area Sub-Scores**, providing recruiters and candidates with an in-depth understanding of result patterns.

The test-taking style measures provide added insight into how the candidate completed the test. Depending on whether the test administered is a supervised or unsupervised version, different forms of test taking style information are provided.

Unsupervised Tests:

Pace – how quickly the individual has responded to the questions

Aptitude – how well the individual has performed on the test

Supervised Tests:

Speed - the proportion of test questions answered in the allocated time

Accuracy - the proportion of correct answers

Caution - the difference between Speed and Accuracy

Saville Assessment online profile charts and feedback reports are designed to be used by a wide audience, including candidates, trained users and line managers, giving more flexibility. Graphic displays are used to ensure quick and straightforward interpretation and feedback.

Typical Feedback Structure

This structure can be used in a variety of contexts, e.g. face-to-face and telephone feedback.

The introduction would include you introducing yourself and your role within the assessment process. It is useful to establish the objective of the discussion, for example to provide feedback after the recruitment process. You could ask the candidate what specifically they would like to get out of the conversation. Ensure you include information such as how long the session will last and who else will see the test results.

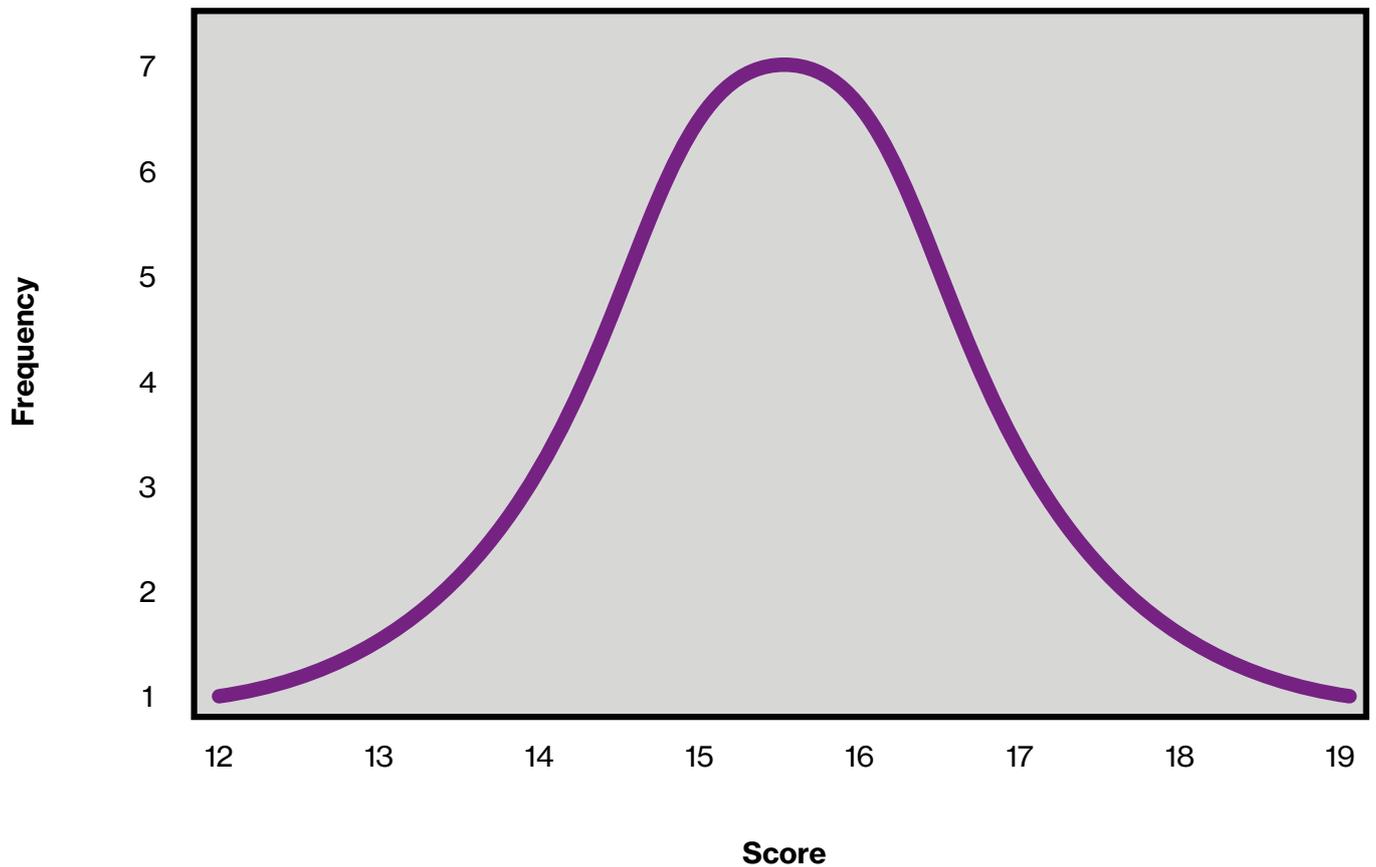
It is worth reminding the candidate of the tests taken, using short and non-technical descriptions, e.g. 'test designed to assess your ability to work with numbers' for a numerical test. Mention the comparison group that candidates' scores are being compared to, so they understand who their scores have been benchmarked against. It is also very useful to ask candidates how they found the tests as it can be useful for you to understand which ones they found easier and which more challenging. This can aid in the discussion of results later on.

Taking each test one by one, describe the candidate's overall score on a test in behavioural terms (e.g. average) and also the proportion of the comparison group they did better than (percentile). Discuss patterns of test-taking style, e.g. slow but accurate and review any differences/patterns in the sub-scores on the profile charts.

Section 7: Test Norms

Frequency Distribution

Score	12	13	14	15	16	17	18	19
Tally								
Count/ Frequency	1	2	3	7	6	3	2	1



Notes:

Exercise: Construct a Frequency Distribution

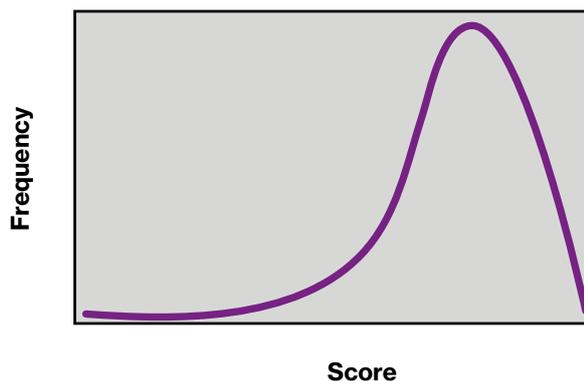
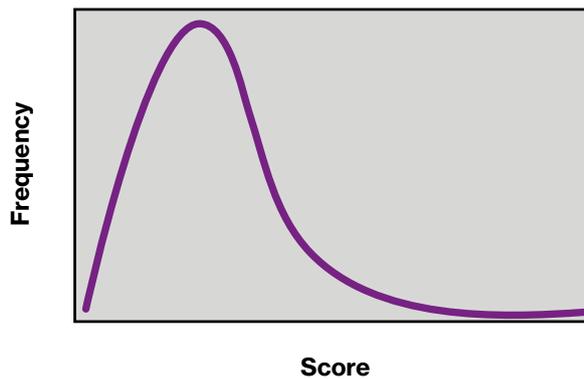
Frequency Table

19, 23, 21, 20, 21, 21, 22, 20, 22, 21

Score	19	20	21	22	23
Tally					
Count/Frequency					



Select the distribution graph that correctly represents the frequency table you've completed



Two Measures of a Normal Distribution

- Where is the middle/average point?
- How spread out is the group?

Measures of Central Tendency

- Mode = most frequently occurring score
- Median = the middle number or midpoint
- Mean (\bar{X}) = the average (total / number of scores)

Notes:

Exercise: Central Tendency

What is the mode, median and mean of the data set?

10, 12, 17, 18, 18, 18, 19, 19, 20, 21, 21, 24

Mode:

Median:

Mean:

Measures of Spread: Range

Equal means on Tests A and B but greater spread of data on Test B

Range is the highest score minus the lowest – but one or two outliers can be misleading

Test A	Test B
54	90
53	80
52	70
51	60
49	40
48	30
47	20
46	10
$\bar{x} = 50$	$\bar{x} = 50$
Range = 8	Range = 80

Notes:

Measures of Spread: Standard Deviation

The standard deviation is a measure of the spread of scores, an average of the difference of scores from the group mean

$$SD = \sqrt{\frac{\sum (X - \bar{X})^2}{N}}$$

SD = the standard deviation

Σ = the sum of

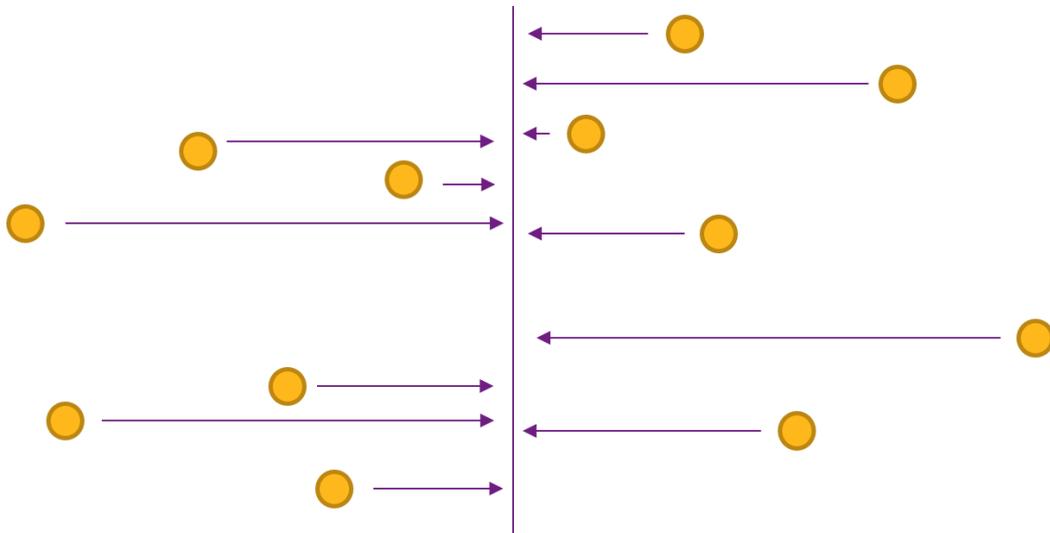
X = a raw score

\bar{X} = the mean score

N = the number in the group

$$\text{Or } SD = \sqrt{\frac{\sum x^2}{N}}$$

x = the deviation score (i.e. raw score minus mean score)



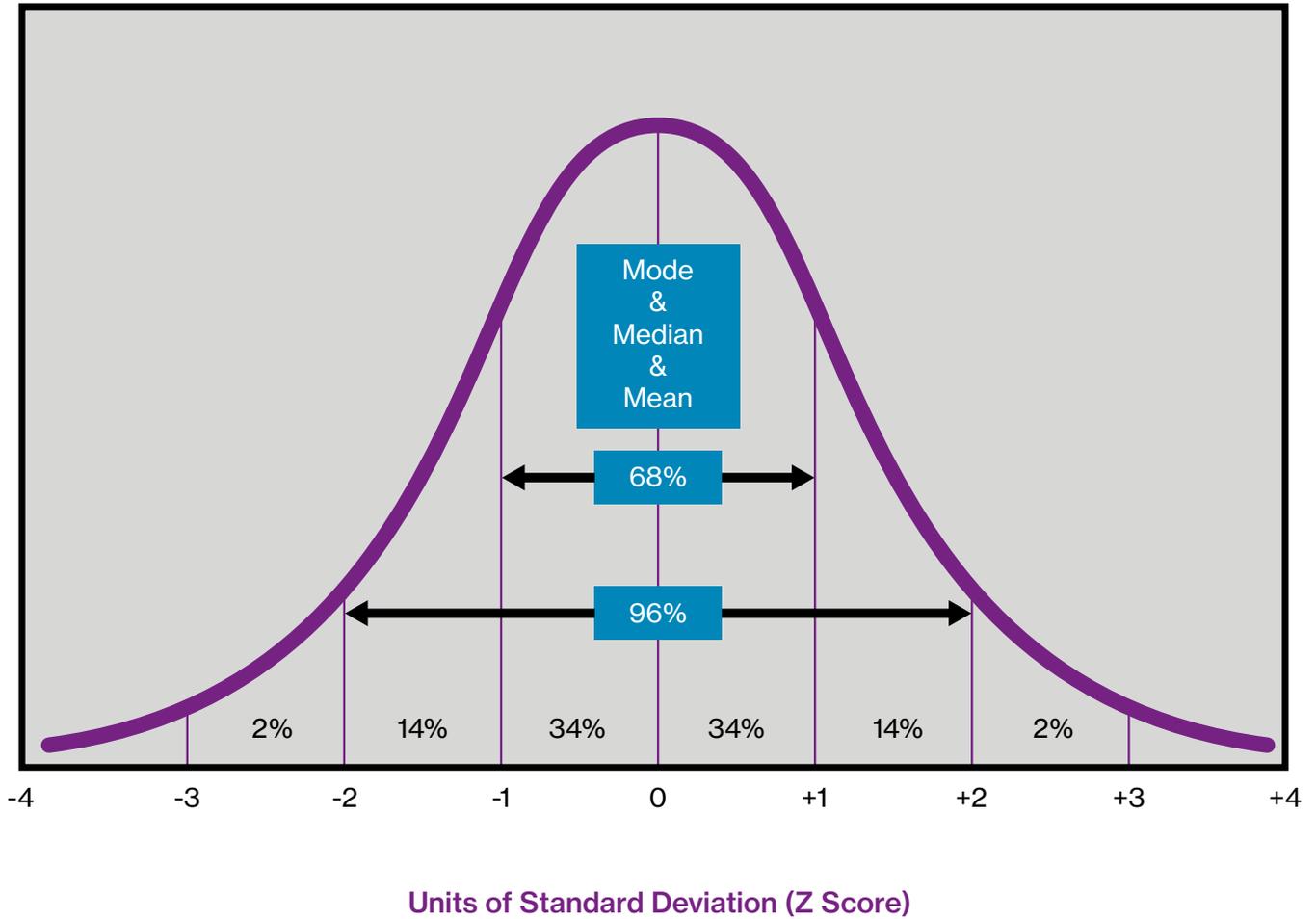
Notes:

Measures of Spread: Standard Deviation

Name	Raw Score X	Subtract Mean to give Deviation Score $X - \bar{X}$	Square the Deviation $(X - \bar{X})^2$
Jones	13		
Grucy	11		
Voison	10		
Bremer	9		
Edge	7		
N =	$\bar{X} = \frac{\sum X}{N} = \underline{\quad} =$		$\sum (X - \bar{X})^2 =$

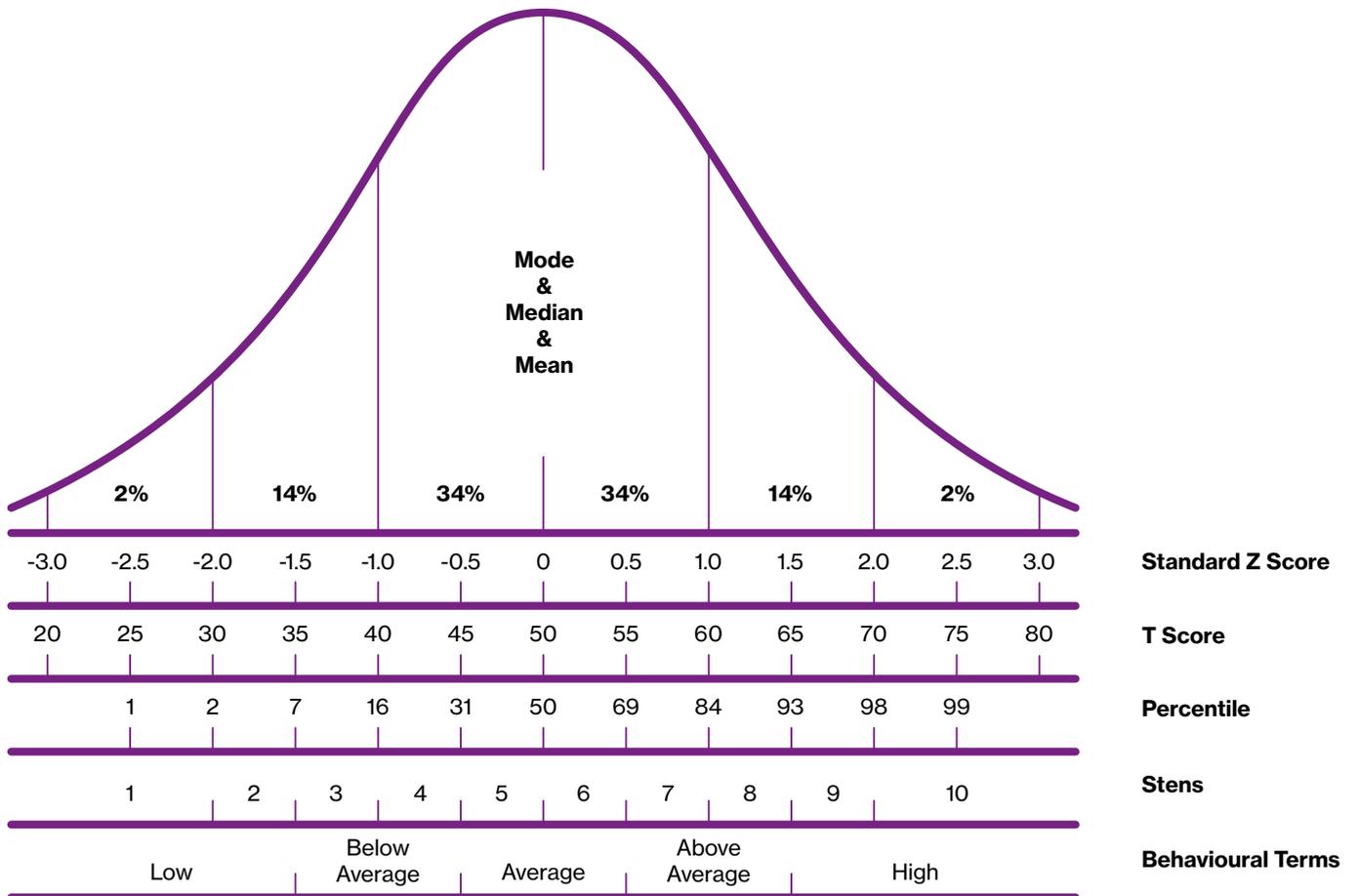
$$SD = \sqrt{\frac{\sum (X - \bar{X})^2}{N}} = \sqrt{\frac{\quad}{\quad}} = \sqrt{\quad} \quad SD = \text{raw scores}$$

SDs and the Normal Distribution



Notes:

Ways of Describing Scores



Notes:

What does SD mean in the real world?



- I am a car manufacturer and want to design a car that fits most men in the UK
- Average Height: 175cm
- SD: 7cm



- So I want to design a car that fits 68% of the population (this is + or - 1SD from average height)
- Which heights does my car fit?



- Now I want to ensure 96% of people fit (this is + or - 2 SD from average height)

Notes:

Selecting Test Norms

Notes:

- Specific vs general benchmark norms
 - Use of specific norms which represent only one protected group, e.g. female only or white only, carries a clear risk of illegal discrimination
- Should take account of:
 - job applied for
 - educational level
 - work experience
 - generalisability / representativeness
- Usually 150+ people
- Up-to-date norms
- Different sampling methods exist, e.g. random, stratified or usage samples

Saville Assessment Aptitude Norms

Geographies

- International
- Regional (e.g. Europe, Africa, Latin America)
- Country (e.g. UK, China, Spain)

Levels/Roles

- Senior Managers and Executives
- Professionals and Managers
- Individual Contributors
- Graduates
- Mixed Occupational Group
 - Mixed Commercial, Customer, Administrative, Operational, Technical
- Apprentices
- Foundation Level
- English as an additional language

Saville Assessment Behavioural Assessment Norms

Geographies

- International
- Regional (e.g. Europe, Asia)
- Country (e.g. UK, France, South Africa)

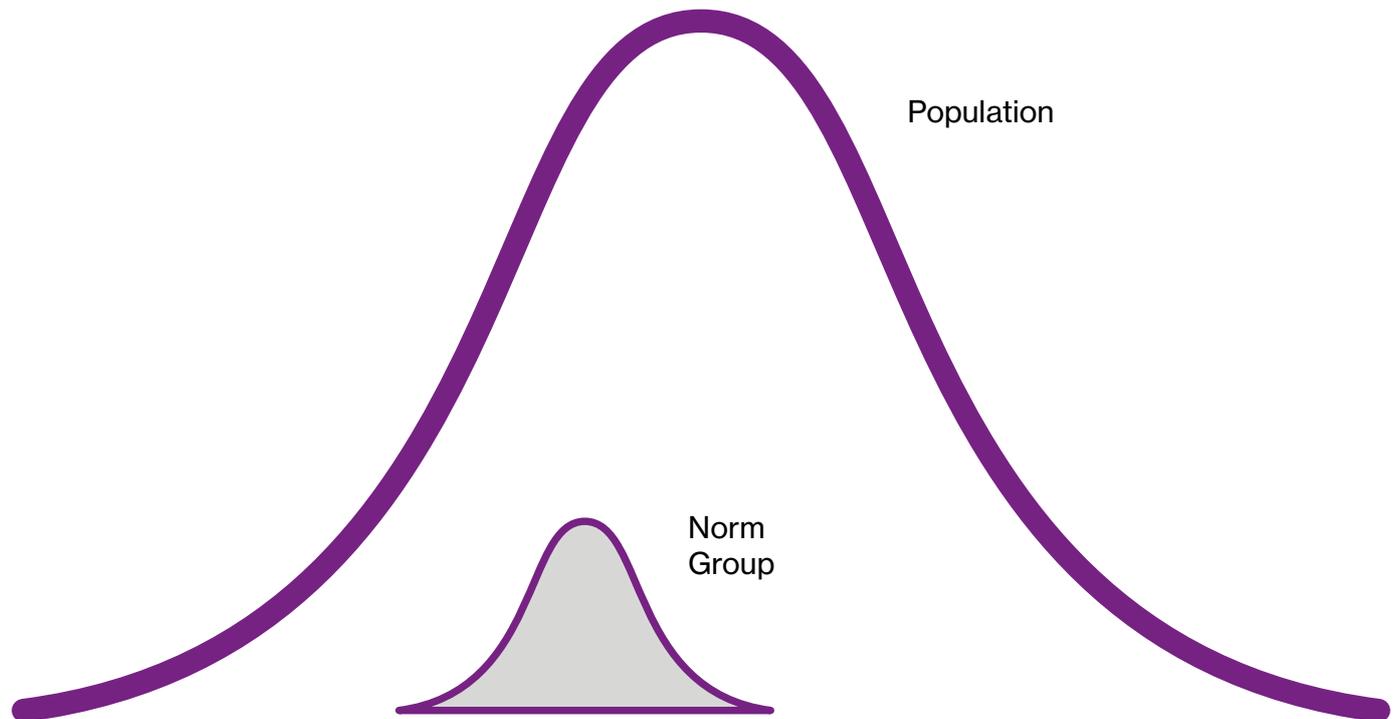
Levels/Roles

- Senior Managers and Executives
- Professionals and Managers
- Individual Contributors
- Graduates
- Mixed Occupational Group

Notes:

How Much Does Size Matter?

- A norm group may not reflect the reality of a whole population
- The standard error of the mean is a calculation that estimates how representative our norm group is of a population
 - As the sample size gets larger, the standard error of the mean reduces and, therefore, the norm group will generally be more representative
- Here the mean and SD for the norm group are smaller than the population – raising possible questions about the representativeness of this sample
- Small norm groups (<150) are more likely to be unrepresentative
- Size matters only to a point – increasing the norm beyond 500 will make little practical difference



Notes:

Case Study: Norm Choice

Q1. Which norm groups would you use for the online unsupervised part of the assessment process you have designed?

Swift Analysis Aptitude or Swift Comprehension Aptitude (delete as appropriate)

- Norm chosen: _____
- Why have you chosen this norm?

Work Strengths

- Norm chosen: _____
- Why have you chosen this norm?

Notes

We have already answered the question of why we use norms and looked at some norm systems often used in tests, such as grades/bands and percentiles. In this section, we'll look at other norm systems we can use. In order to do this, there's a little more to learn about the normal distribution.

The Normal Distribution

There are two key measures of a normal distribution; where the middle is and how spread out the data is. By understanding these measures and how they impact on the shape of the curve, we can start to understand things like how easy or difficult a test is.

Where is the middle?

This relates to the average of the group scores. There are several ways to identify the average of a set of scores:

The Mode - The mode is the most frequent score in the set of scores. Occasionally you may have more than one mode in a set of scores.

The Median - The median is identified by lining up scores in order and finding the middle number. It is the number which has 50% of the scores above it and 50% of the scores below. If there are even numbers of scores, then calculate the number which would be half way between the two middle numbers.

The Mean (\bar{X}) - The arithmetic mean is often referred to as the average of the set of scores. The mean is calculated by adding all of the scores (**X**) in the group to find the total (**ΣX**) and then dividing the total by the number of people in the group (**N**). In statistics it can be calculated using a formula where the statistical signs have the following meaning:

Formula: Mean (\bar{X}) =
$$\frac{\Sigma X}{N}$$

If a set of scores is normally distributed then the mode, median and mean all have the same value and coincide with the peak of the curve. If the distribution is other than normal, i.e. skewed, then the mode, median and mean may assume different values to each other.

How spread out is the group?

The Range

One obvious measure of the spread of the group is the range. The difference in scores between the candidate with the highest score and the candidate with the lowest score is called the range.

The range can add significant detail to the picture. Consider the scores shown in the table here. Both Test A and Test B show a mean score of 50, however Test B shows much greater variation in the scores of the group (the lowest is 10 and the highest is 90) compared to Test A (where the lowest score is 46 and the highest 54).

The range can, however, be misleading if the set of scores you are looking at contains one or two outliers (that is one or two candidates who have either scored very low or very high). This can give you a high range score, but does not give you a good general idea of the actual spread of scores for the whole group.

Test A	Test B
54	90
53	80
52	70
51	60
49	40
48	30
47	20
46	10
$\bar{X} = 50$	$\bar{X} = 50$
Range = 8	Range = 80

The Standard Deviation

The Standard Deviation (SD) is a statistic which tells us about the spread of scores around the average or mean and gives us a more robust measure of what is termed 'dispersion' of scores. This is essentially the average of the spread of scores around the mean of those scores. The SD tells you on average how far away scores are from the average score – it is a measure of the variability of the scores in that group. Thus, a group of scores can exhibit a high degree of **variability** (high SD) or a low degree of variability (low SD), indicating a more homogenous group.

The formula looks complicated at first sight, but on closer inspection, the calculations are not too complex. It is best to view the formula as a recipe, with several steps that need to be taken.

There are six simple steps to calculate the Standard Deviation:

1. Calculate the mean of the raw scores
2. Subtract the mean from each raw score to obtain the deviation score $X - \bar{X}$
3. Square the deviations to obtain the $(X - \bar{X})^2$ values
4. Sum [Σ] the squared deviations to obtain the $\Sigma(X - \bar{X})^2$
5. Divide the summed squared deviations by the group number [N]
6. Find the square root of this value – this is the SD

$$SD = \sqrt{\frac{\Sigma (X - \bar{X})^2}{N}}$$

$$\text{Or } SD = \sqrt{\frac{\Sigma x^2}{N}}$$

Example:

In this example, the mean of the set of five scores is 10 raw scores and the standard deviation is 2 raw scores. That means that on average, the scores in this group differ from the overall mean (10) by 2 raw scores. This is a descriptive statistic, giving us a way of describing how variable the scores in the group are. We can say that overall, the variability in this group is quite low.

Name	Raw Score X	Subtract Mean to give Deviation Score $X - \bar{X}$	Square the Deviation $(X - \bar{X})^2$
Jones	13	+3	+9
Grucy	11	+1	+1
Voison	10	0	0
Bremer	9	-1	+1
Edge	7	-3	+9
N = 5	$\bar{X} = \frac{\sum X}{N} = \frac{50}{5} = 10$		$\sum (X - \bar{X})^2 = 20$

$$SD = \sqrt{\frac{\sum (X - \bar{X})^2}{N}} = \sqrt{\frac{20}{5}} = \sqrt{4} \quad SD = 2 \text{ raw scores}$$

In this example, with a group mean of 10 and standard deviation of 2, how many raw scores would a candidate need to score to do better than one standard deviation above average?

The answer is 12.

If someone scored 16 on the test in the example, how many standard deviations is this score either above or below average? The answer is 3 standard deviations or 3 units of standard deviation.

So standard deviations are the average departure of individuals' test scores from the group average. Units of standard deviation can be plotted onto a normal distribution of scores, above and below the average score. You can expect to see certain percentages of individuals' scores falling between defined areas under the normal distribution curve:

- Between +1 and -1 standard deviations (i.e. one above average and one below average) you can expect to find 68% of people's scores.
- Between + 2 and – 2 standard deviations, you can expect to see 96% of people's scores. Between + 3 and – 3 standard deviations, you can expect to see 99% of people's scores.

These are statistical facts of life associated with the normal distribution.

For example, if the average height of men is 5' 9" and the standard deviation is 3", 68% of men are likely to be between 5' 6" and 6' 0". 96% of men are likely to be between 5' 3" and 6' 3". If a man is 5' 6" tall, how many units of standard deviation from the mean height of all men is he? The answer is 1 standard deviation below average (-1 SD). If a man is 6'3" tall, how many units of standard deviation from the mean height of all men is he? The answer is 2 units of standard deviation above average (+2 SDs).

Other Scoring Systems

Z Scores

By looking at units of standard deviation, we have been using a new scoring system without even knowing it! We have been using the Z score system. Z scores are a vital concept in psychometrics and form the basis of all standard score norms.

The Z score scale represents the number of standard deviation units a raw score is above or below the group mean. Z scores are equal units of measurement and (unlike percentiles) you can add Z scores from different tests to produce overall composite scores. To calculate the Z score you need to know the group mean and the group standard deviation. The Z score is calculated using the formula presented here.

$$Z = \frac{X - \bar{X}}{SD}$$

Example: Jim scores 80 on a test where the Standard Deviation (SD) is 10 and the Mean is 50. What is Jim's Z score?

$$Z = \frac{80 - 50}{10} = \frac{30}{10} = +3$$

This means that a raw score (X) of 80 on this test results in a Z score of +3, that is, 3 standard deviations above average (three lots of 10).

Z score values are simply the number of units of standard deviation that a raw score is away from the average. Typically, the Z score scale runs from approximately +3 to -3. The average of a Z score scale is 0 and the standard deviation of a Z score scale is 1.

There are some disadvantages in using Z scores to describe people's scores on tests – they are decimals and they also have positive and negative signs. This makes them difficult to interpret from the test user, line manager and candidate point of view and, whilst it is possible to produce composite scores, it is sometimes difficult to do other mathematical calculations with them.

T Scores

To overcome the issues created by having a scale that runs through from positive to negative, we can simply transform the Z score into a T Score (Transformed Score). The formula is presented here:

$$Z = \frac{T - 50}{10}$$

The T Score scale has an average of 50 and standard deviation of 10. Whilst T scores are widely used in the interpretation and analysis of ability test scores, they aren't particularly candidate friendly and so you will rarely see them on reports.

Sten scores

The Sten scale (or standard ten scale), is another way of describing scores. Stens tend to be used when measuring personality, style or motivation, because less detailed differentiation is required for these attributes. The Sten scale divides scores into 10 categories, with 1 being low and 10 being high. The mean of the Sten scale is 5.5 and the standard deviation is 2. When describing Stens, it is customary to round up the Sten to the nearest whole number. The disadvantages of using Stens are that they are broad bands, like grades, and you cannot make fine-tuned judgements about differences between scores. However, converting ability test scores to Stens can be useful in assessment or development centre situations, where other assessment exercises have been scored or rated using a 1 to 10 scale.

To calculate a Sten from a given Z score, you multiply the Z score by 2 and add 5.5.

Norm Groups

A norm group is the sample group against which a candidate's scores are compared. A norm group can be regarded as a sample, from which a set of scores have been gathered to provide a representation of the population it is intended to represent (e.g. managers, graduates, call centre staff or the general population).

Norm groups should be up to date and in order to have statistical significance, should be based on a group of 100 + people. The extent to which the sample is representative of the population depends greatly on the size of the sample. Larger samples are more likely to be representative. For example, a norm group of a thousand Managers is more likely to be a reasonable representation of all Managers than a group of 50. The standard error of the mean (SEmean) allows us to estimate the distance of our sample mean from the population mean. It can be calculated using the mean of the sample, the SD and the sample size. As the sample size gets larger, the SEmean will reduce.

A large group, however, is not guaranteed to be representative and attention should be paid to the sampling method that has been used and whether there is an appropriate spread of people in the group. There are a number of different ways of collecting samples and developing norms. The main sampling methods include random, stratified and usage sampling.

In random sampling, you randomly select people to include in your sample, typically with the entire population having equal chance of being selected – this can be challenging to achieve when your population is large and broad, but can reduce sampling bias.

Stratified sampling is where you purposely select a sample that is representative of your intended population (e.g. if your population consists of 50% males and 50% females, you would ensure the same proportions in your sample).

However, even with a sophisticated methodology, there is still the challenge in assessment of the motivations of the candidates. Therefore, the final data from random or stratified samples are likely to be different from realistic live usage samples.

In usage sampling, you include those in your sample who have previously completed a test in a real application. Usage samples are often used to collect norm samples; an advantage of this is that those within the sample are usually realistically motivated to complete the test(s). Many usage samples and therefore norms are merely opportunist in nature, however, this is not the Saville approach.

We combine the advantages of using motivated usage data with the techniques of careful stratified sampling to create norm groups which are both representative and based on realistic test completion data. For example, Saville prevent over-sampling of particular groups when creating realistic usage norms by reducing the number of individuals selected from a particular country or client. This leads to smaller but better norm groups. Because we know that, statistically, beyond a few hundred people, sample size has relatively less impact than sample composition on norm quality, our best practice recommendation is to use realistic usage samples which are cut down to a smaller, more representative norm group.

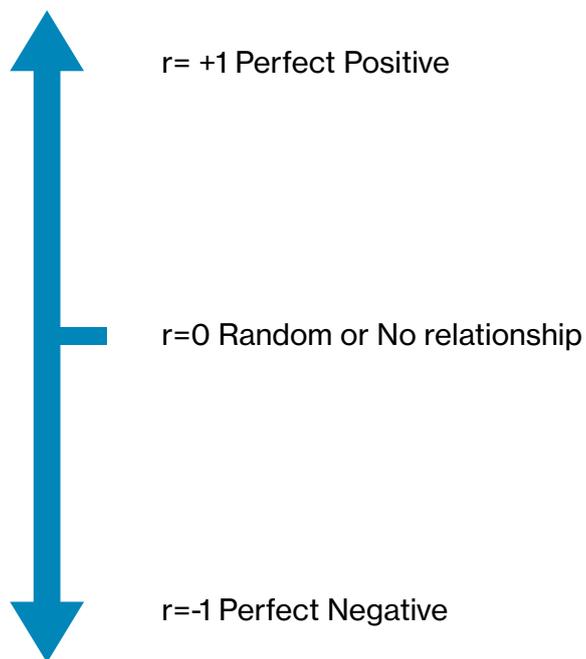
Choosing Norm Groups

The group used to establish an appropriate norm group should always take into account the job being applied for. It would be appropriate to use a norm group of graduates for graduates entering the organisation. It would not be appropriate to compare graduates' scores on a numerical test against a group of 16-year olds, nor would it be appropriate to compare them to a group of experienced Chartered Accountants. It is therefore important to take into consideration things like educational level and work experience in order to ensure that your norm group is representative of your candidate population.

Section 8: Correlation

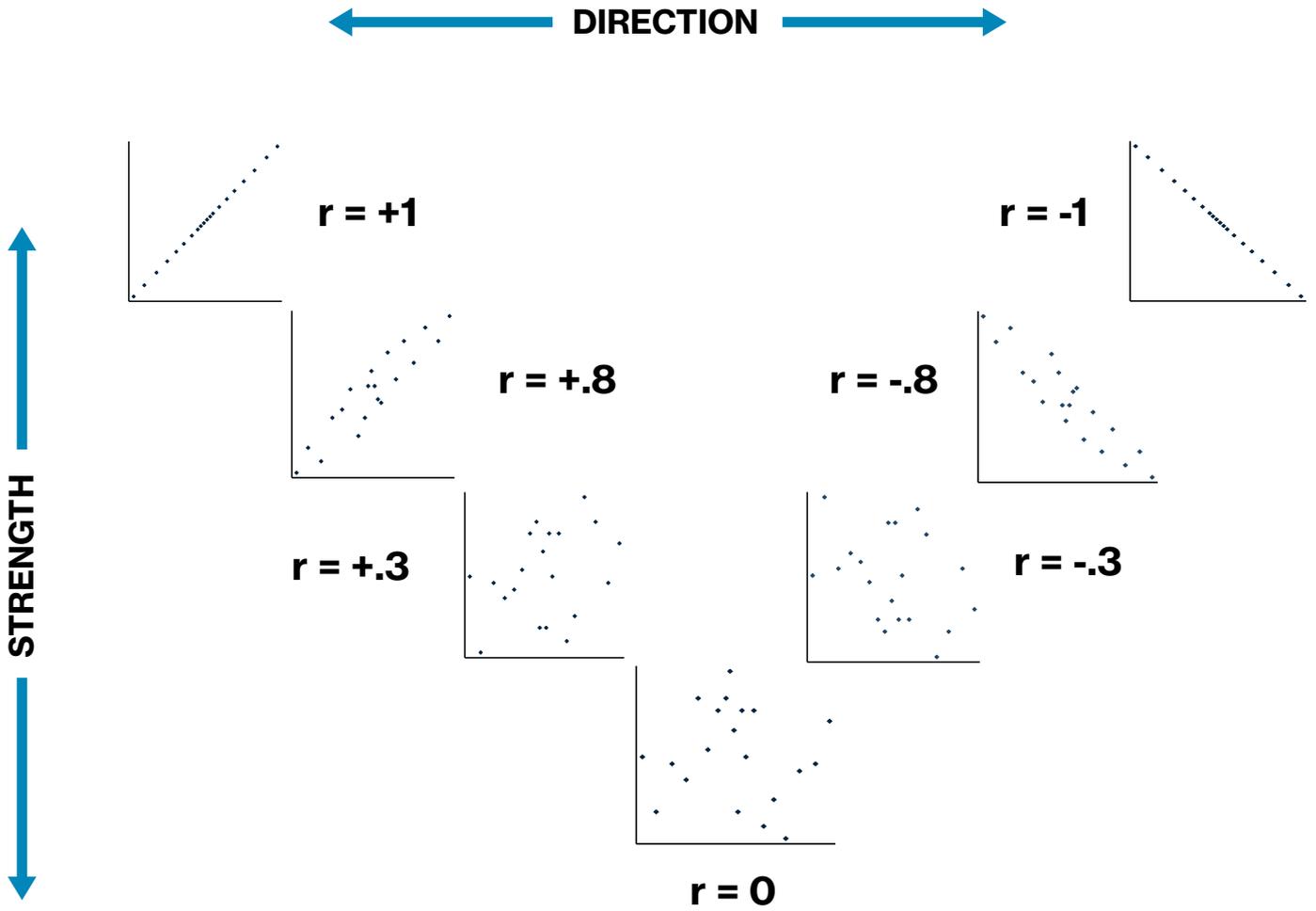
What is a Correlation?

- Correlation is a statistical technique for establishing whether there is a relationship between two things
- The degree of relationship is expressed using a correlation coefficient (r), the further away from 0 the stronger the relationship
- Correlation is commonly used within research to establish reliability and validity



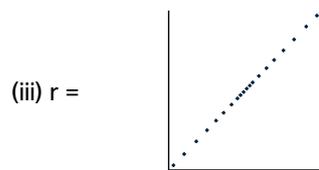
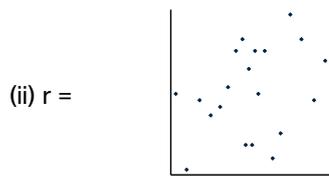
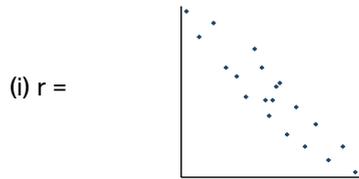
Notes:

Interpreting Correlation Strength



Exercise: Scattergrams

For each of the scattergrams below, indicate approximately what degree of correlation is shown using the correlation scale.



(i)

(ii)

(iii)

(iv)

Maximising or Minimising Correlations

No time delay between measures	Maximises
Many influencing factors	Minimises
Poor measures of variables	Minimises
A strong underlying relationship between two variables	Maximises

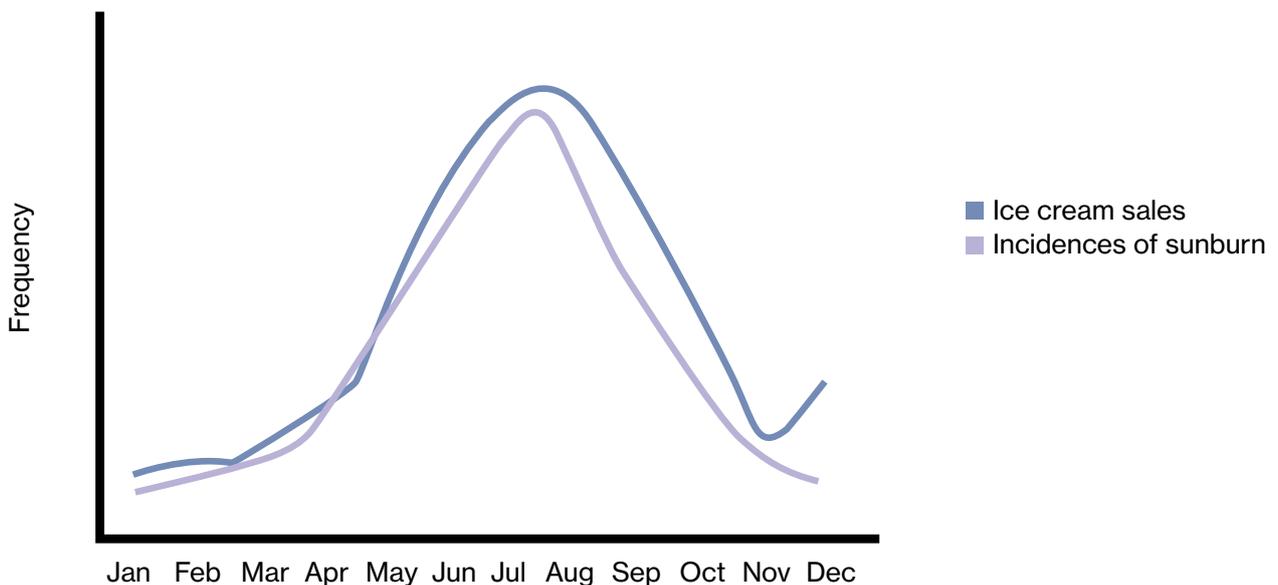
Interpreting a Correlation

Are we confident?

- Statistical significance gives you a level of confidence in the result
- Lack of statistical significance indicates that the result is likely to be a chance finding

Does one thing really lead to another?

Ice cream causes sunburn



Notes

The Importance of Correlations

Correlation is used to test **reliability** and **validity**; two incredibly important concepts when it comes to testing.

Reliability is about the precision or accuracy of measurement of a test or assessment.

Validity is about enabling the best possible decision. If a test is valid, it measures (or predicts) what it is designed to measure (or predict). If a test is valid, then decisions based upon it will lead to you selecting people who will perform better on the job than others you may have picked from the group.

What is Correlation?

The correlation coefficient was first conceived by early psychologist Sir Francis Galton, who also discovered isobars and fingerprints. Subsequently refined by statistician Karl Pearson, the main coefficient became known as **Pearson Product Moment Correlation Coefficient** and is universally given the letter 'r'.

The development of correlation is widely recognised as one of the most important mathematical achievements in the last 200 years and is used now in psychology, medicine, engineering, physics, market research and economics.

Correlation coefficients show how two variables relate to each other. To run a correlation on a group you need to be able to measure two things about each person in the group and then you can calculate how they are related.

Correlation answers questions about a group of people, e.g. 'Are taller people heavier in this group?' Relationships can be assessed both as a matter of direction and strength.

The more you eat, the more you weigh – a **positive relationship** as if higher on one variable also likely to be higher on the other variable. The more energy you use, the less you weigh – a **negative relationship** as when one variable is higher the more likely it is to be lower on the other variable. Correlations vary in strength – from being perfect where you can exactly predict scores on one variable from another to being very weak.

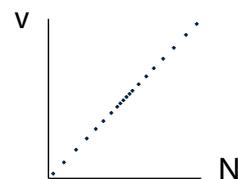
Scattergrams

Perfect Positive

The scattergram is the first and easiest way to identify whether two variables or sets of data are related. The verbal and numerical ability test scores for a group of five candidates have been collected and plotted on the scattergram below.

Each diamond represents one candidate and the point at which their scores intersect. We can see that high scores on the verbal test are associated with high scores on the numerical test. By the same token, low scores on the verbal test are associated with low scores on the numerical test.

The result is a straight line running from the bottom left to the top right of the scattergram – this indicates a positive correlation. As the line is also straight, it shows a perfect positive correlation. We could even estimate the expected numerical test score, knowing just the verbal test score, e.g. a candidate gaining a verbal test score of 30 will achieve a numerical test score of 24. This is known as a perfect positive correlation; scores are increasing in the same direction and we have the straight line on the scattergram. If we were to calculate the exact **Pearson's Product Moment Correlation Coefficient**, the result would be $r_{xy} = +1$. We could save money by just giving candidates one of the tests as we could predict their scores on the second.

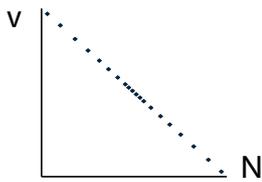


Perfect Negative

In the scattergram below we can see that high scores on the verbal test (are associated with low scores on the numerical test. By the same token, low scores on the verbal test are associated with high scores on the numerical test.

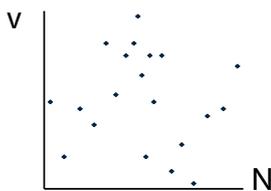
The result is a straight line running from the top left to the bottom right of the scattergram – this indicates a negative correlation. As the line is also straight, it shows a perfect negative correlation. We could even estimate the expected numerical test score, knowing just the verbal test score, e.g. a candidate gaining a verbal test score of 30 will achieve a numerical test score of 0. In this situation, scores are going in different directions. If we were to calculate the exact correlation, the result would be $r_{xy} = -1$. As with a perfect positive, we could save money by just giving candidates one of the tests, as we could predict their scores on the second.

Perfect correlations are very rare between two different things measured, but do result from converting a measure of the same thing on one scale to be on a different scale. Height in inches with height in cm correlates perfectly when converted (i.e. by multiplying height in inches by 2.54).



No Relationship

In this third example, we can see that there is no straight line on the scattergram, indicating no discernible relationship between the verbal and the numerical test scores. We cannot predict one score from knowledge of the other. In this instance, we have a zero or no correlation, indicating no relationship between the two variables or sets of data. Scores on the verbal test seem to be independent of scores on the numerical test for this group of candidates. The correlation coefficient is near zero - $r_{xy} = 0$.



Strength of Relationship

By looking at the scatter of scores we can look at the strength of the relationship and the direction of the relationship between two variables without doing any statistical calculations; you are simply 'eyeballing' the data.

How close the data is to a straight line tells you about the **strength** of the relationship (how close to 1), and whether the line is angled to the left or to the right tells you about the **direction** of the relationship (positive or negative). This is an easy way to investigate the relationship between two variables but is very time-consuming when it involves large quantities of data, and really only gives a visual indication rather than anything more specific.

Statistical Significance

Statistical significance tests are concerned with whether a correlation you have calculated is non-zero, i.e. whether there is a relationship between the two variables.

To ascertain whether a significant relationship exists, you need the following information:

1. The correlation you have calculated (e.g. $r = .32$).
2. Whether you are only looking for a correlation to be positive (one-tailed test). For instance, you might only want to consider whether a test positively correlates with managerial rating of work performance. If you are only looking for correlation to be negative, this would also be a one-tailed test. If you are interested in whether a correlation is positive or negative, you are running a two-tailed significance test.
3. Calculation of the degrees of freedom (df). If your hypothesis is two-tailed, degrees of freedom = number of 'cases' minus 2. For example, if you have 42 participants, then there are 40 degrees of freedom for your correlation coefficient. If your hypothesis is one-tailed, degrees of freedom = number of 'cases' minus 1.
4. What probability you require to establish that your correlation is statistically significant (non-zero). The most common probability is $p < 0.05$ (where less than 5 in 100 or 1 in 20 occasions our results could be down to chance factors).

Points on Correlation

- A significant correlation coefficient does not indicate cause and effect. It is common that a correlation may be caused by an external factor which causes both variables to change systematically.
- Different populations may have different correlations. Age and verbal reasoning in a population of children is not the same as in a population of adults.
- The smaller the sample, the greater the sampling error and the greater the chance of obtaining a misleading result. In practice, this means that looking for small or moderate correlations (which characterise much of personnel selection) with samples of less than 100 does not give results with much meaning. To get a result from one sample that has small enough confidence intervals for meaningful interpretation ideally requires a sample of 200 or even higher, although small samples can be a useful contribution to studies which integrate the results from many studies, i.e. meta-analysis.
- Restriction of range in samples results in correlations being smaller than in the wider population. If you exclude the bottom performers on the test you will in practice only see half the correlation on the scattergram when you plot against the performers on the job (i.e. you cannot collect the data on the low performers on the test as you have not recruited them).

Notes:

Section 9: Reliability

What is Reliability?

- Nothing can be measured with absolute accuracy
- A test's reliability concerns the **precision** and **consistency** of measurement
- **Classical Test Theory (CTT)** assumes that:
 - Observed score = true score + margin of error
- The more error in test scores, the lower the reliability
- Generally the longer the test, the greater the reliability
 - NB: Use total scores on Swift tests for decision making
- Reliability is a prerequisite for validity

Notes:

Possible Sources of Error

Candidate

- Feeling unwell
- Misinterpreting instructions
- Severe test anxiety

Administrator

- Using an unreliable test
- Poor candidate briefing
- Misinterpreting test results

Test Developer

- Ambiguous items
- Items measuring the wrong thing
- Poor instructions

Methods of Assessing Reliability

Test-Retest – administering the same test twice

- + Gives indication that attribute is stable
- Candidates not willing to do it twice

Alternate Form – administering two parallel forms

- + Shows developer is clear/consistent on what is measured
- Has the expense of developing two forms

Internal Consistency – administering only one test once

- + Easy to do as only requires one test on one occasion
- Can be misleadingly high with repetitive item content

Notes:

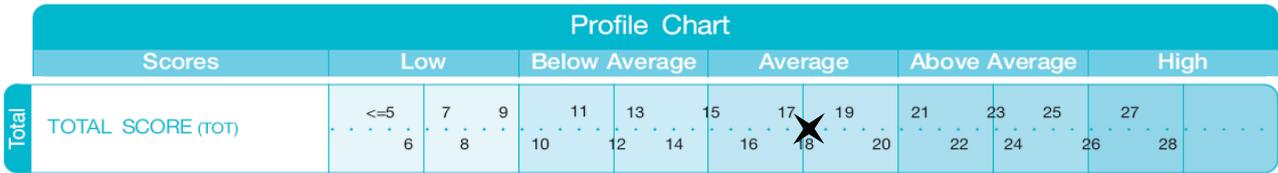
Professional Aptitudes Reliability

Reliability of Saville Assessment Professional Aptitudes V/N/D

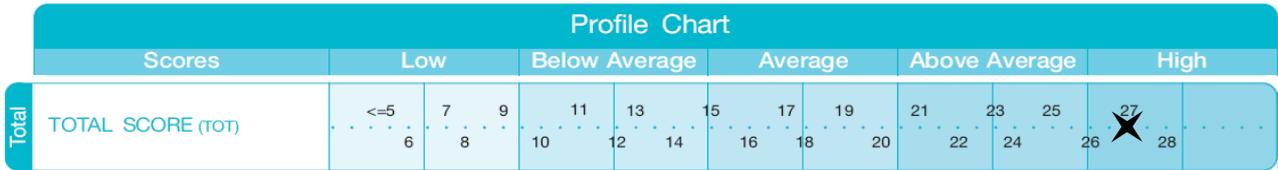
	Mean Maximum 28	SD	r Internal Consistency Reliability (N=300)	r Test Retest Reliability (N=120)	r Alternate Form Reliability (N=135 V&N; N=109 D)
Professional Verbal Analysis	17.75	5.51	.84	.76	.71
Professional Numerical Analysis	13.94	5.35	.87	.72	.70
Professional Diagrammatic Analysis	18.26	4.48	.80	.71	.69
Average	16.63	5.11	.84	.73	.70

How Different are These Scores?

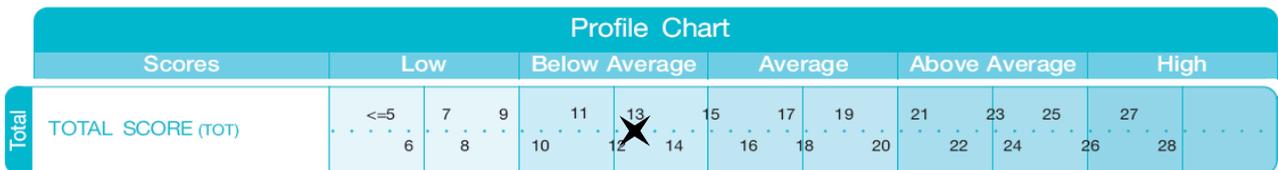
Candidate A



Candidate B



Candidate C



Estimating the Margin of Error

- No workplace assessment score is perfectly reliable or absolutely precise
 - Although aptitude tests are known to be one of the most reliable and precise of all assessment measures
- The more error in an assessment, the less precise the result
- The **Standard Error of Measurement (SEm)** estimates the margin of error around test scores
- The greater the reliability, the smaller the band of error around a test score

The Standard Error of Measurement

The Standard Error of Measurement (SEm) is the Standard Deviation of a group multiplied by the square root of one, minus the reliability coefficient for the test.

$$\text{SEm} = \text{SD} \sqrt{1 - r_{tt}}$$

SEm = Standard error of measurement

SD = Standard deviation of scores for the reliability sample

r_{tt} = the reliability coefficient (test-retest, alternate form or internal consistency)

Notes:

Calculating the SEM

Total Raw Score (X) = 10 on Verbal Analysis

What is the expected margin of error for this score?

If: SD = 4 Raw Scores, $r_{tt} = 0.75$

$$\begin{aligned}\mathbf{SEm} &= \mathbf{SD} \sqrt{\mathbf{1} - \mathbf{r}_{tt}} \\ &= \mathbf{4} \sqrt{\mathbf{1} - \mathbf{0.75}} \\ &= \mathbf{4} \sqrt{\mathbf{0.25}} \\ &= \mathbf{4 \times 0.5} \\ &= \mathbf{2 \text{ Raw Scores}}\end{aligned}$$

Notes:

Applying the SEM: Error Bands

What is the band of error around a score of 10?

+/- 1 SEM = 68% probability of band capturing true score*

1 SEM we calculated as equal to 2

So for a score of 10 +/- 2 gives band of error of 8 to 12

+/- 2 SEM = 96% probability of band capturing true score

2 SEMs would be $2 \times 2 = 4$

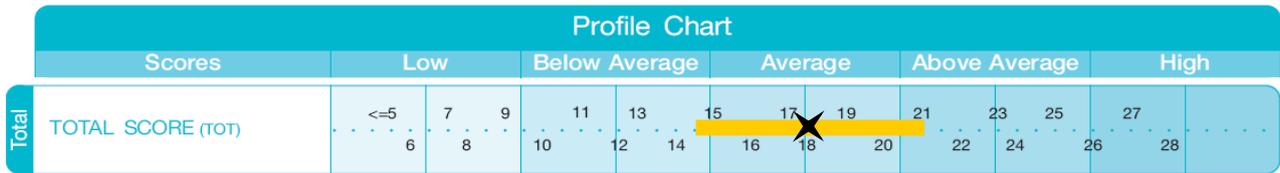
So for a score of 10 +/- 4 gives band of error of 6 to 14

*We use the 68% or 1 SEM band of error on our reports

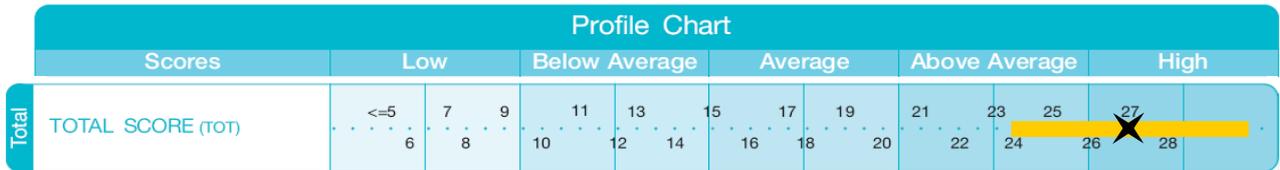
Notes:

How Different are These Scores?

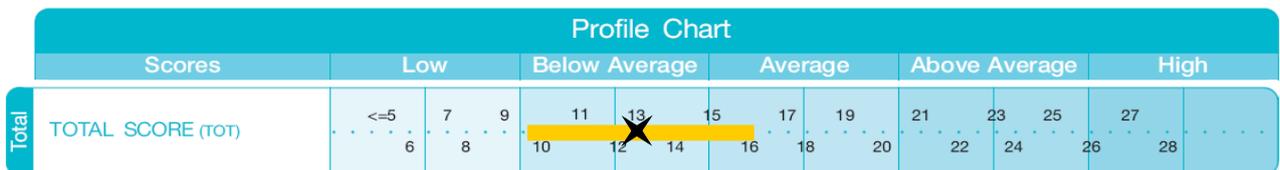
Candidate A



Candidate B



Candidate C



Case Study: Reliability

The hiring manager isn't convinced by your proposed assessment process. Specifically, they have concerns around the use of aptitude tests. They are not experienced in psychometrics themselves and argue that low test scores could just be reflective of a candidate having a bad day.

You must present a counter argument to the hiring manager that highlights the consistency and precision of your Swift test in the online unsupervised part of your assessment process.

Q1. Use the technical documentation to find reliability estimates for the Swift test chosen.

Swift Analysis Aptitude or Swift Comprehension Aptitude

- Reliability:

Q2. Write some notes on how you would present a counter argument to the manager.

Notes

Definition of Test Reliability

Reliability is fundamental to measurement and concerns how precise and error-free a tool is in measuring desired constructs. Any instrument that measures something in the real world needs to have a level of precision or accuracy (i.e. weighing scales, a digital clock, a light meter in a camera).

The greater the reliability or precision, the greater the chance that it will allow for valid decision-making. If your watch could be an hour out either way any morning it makes it difficult to decide when to leave for work to get there on time.

Test publishers aim to produce tests that are robust enough to reflect people's abilities, rather than reflecting the conditions on the day of testing. Test publishers calculate estimates of test reliability when developing tests; it is not the responsibility of test users.

The greater the reliability of a test, the more likely it is to show better validity. However, good reliability is no guarantee that a test will help you make better decisions, i.e. a highly accurate mechanical test is unlikely to make valid decisions about who would be the best accountant in the company.

Classical Test Theory states that: A person's obtained raw score = true score + error.

There are two main criticisms of reliability theory: (1) the assumption that error is always random; (2) both measurement error and the reliability coefficient can vary, thus neither is a stable characteristic of a test.

Sources of Test Error

Error can creep into the testing process from a variety of sources: the test administrator, the environment in which testing has taken place, the candidate and also the test developer or publisher. This puts a premium on test administrators running professional and standardised test administration sessions to minimise error and maximise reliability. Use of a stopwatch and 'backup' source of timing will ensure that tests are run to the correct time, every time.

Ensuring that candidates are given the correct administration instructions will reduce error. This puts a premium on reading administration instructions accurately and clearly. Coaching candidates through practice

questions to ensure they correctly understand the task ahead is also crucial. From an environmental perspective, ensure that the testing room is well lit, at an appropriate temperature and is quiet enough to allow candidates to concentrate. Putting a 'do not disturb' sign on the door should also reduce distractions and avoid introducing further error.

Error can also come from candidates experiencing temporary states such as boredom, fatigue, anxiety or illness. A well-run test session should aim to inform, motivate and relax candidates. Test publishers aim to reduce the ambiguity of items to ensure that they are accurate assessments of performance of what they are assessing. Items on ability assessments can vary greatly between tests and test publishers in how effective they are. Items can be ambiguous so that there is no clear answer or that more than one answer could be correct.

Types of Reliability

Test-Retest reliability refers to the stability of a measure over time. It is calculated by correlating scores on a measure completed by the same group of people at two points in time.

Alternate or Parallel Form Reliability refers to the consistency between two versions of the same measure. This is the correlation between the results for the same group of people who complete two versions of the test.

Internal Consistency Reliability relates to the internal correlations of the components of the measure, for example the relationship between the different scales within an assessment.

While all forms of reliability are important, internal consistency is often the most practical and accessible form of reliability, which can be more readily calculated in large samples.

The accepted benchmark level for test reliability is $r = +0.70$.

Reliability figures for the Diagrammatic Analysis tests tend to be slightly lower than those for Verbal and Numerical Analysis but average at a satisfactory level of .70. The format of this type of test is somewhat unfamiliar at first so that substantial 'within test' and 'across tests' practice effects can occur. It is therefore of great importance

to make Preparation Guides available to all candidates, especially when using the short Swift Analysis Aptitude measure.

Margin of Error/Error of Measurement

As discussed previously, nothing is ever completely accurate and there is always some error of measurement that needs to be taken into account. The key questions are: (1) How much fluctuation in a person's test score should we allow? (2) Are two people's test scores really different from each other?

You may recall that when you were profiling up your own test scores using the Saville Assessment profile charts, you drew a band of error in T scores either side of your total score, test-taking style scores and item type sub-scores. This band shows visually the amount of error to draw around a person's test score. If a person was to complete the same test over several different occasions, their test score will form a **mini normal distribution**; this is in contrast to the larger normal distribution, which illustrates single test scores for many people.

The less reliable a test is, the bigger the band of error that needs to be drawn around a person's test score. The more reliable a test is, the smaller the band of error that needs to be drawn around a person's test scores.

The mini normal distribution represents the Standard Error of Measurement (SEm) for that particular test. Test publishers provide details of the SEm for their tests, either on the profile charts (which is what we do for Saville Assessment tests) or in the user manuals for other test publishers. For the purposes of this course, you also need to know how to calculate an SEm, which is what we'll come on to next.

SEm Equation

$$SEm = SD \sqrt{1 - r_{tt}}$$

To calculate the SEm, you need to know the Standard Deviation of scores for a particular group on a test, and you also need to know the reliability of the test, expressed as a correlation coefficient (r).

The SEm can be calculated as a Raw Score, a Z score, a T Score, a Sten or an IQ, but not as a percentile. The units of

the SEm depend on the units of the Standard Deviation. For example, if the Standard Deviation is a Raw Score, the SEm is also expressed as a Raw Score. The SEm is applied as a plus and minus figure around a candidate's obtained score. SEm figures should be found in the relevant test manual.

Calculating the SEm

As an example, let's say that someone achieved a raw score of 10 on Professional Verbal Analysis and we know that the Standard Deviation of the group in raw scores is 4. The test-retest reliability estimate of Professional Verbal Analysis is 0.75.

To calculate the SEm, follow these steps:

1. Insert the Standard Deviation and Reliability figures into the equation.
2. Subtract the Reliability estimate from 1. e.g. (1 – 0.75).
3. Using your calculator, find out the square root of the answer to step 2.
4. Multiply the answer to step 2 by the Standard Deviation (4 x 0.50)
5. This is your SEm, i.e. 2 Raw Scores

Applying the SEm

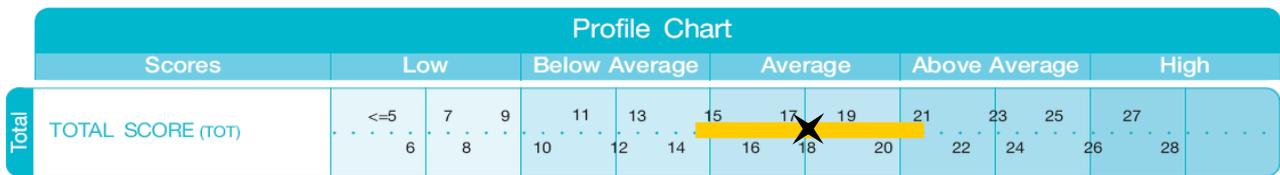
The SEm works on the same principles as the Standard Deviation in terms of confidence intervals. We can be confident that 68% of the time, a person's score will lie within plus or minus 1 Standard Error of Measurement of their score. We can be confident that 96% of the time, a person's score will lie within plus or minus 2 Standard Errors of Measurement of their score.

In our example, the band of error at the 68% level is 8 to 12 and the band of error at the 96% level is 6 to 14.

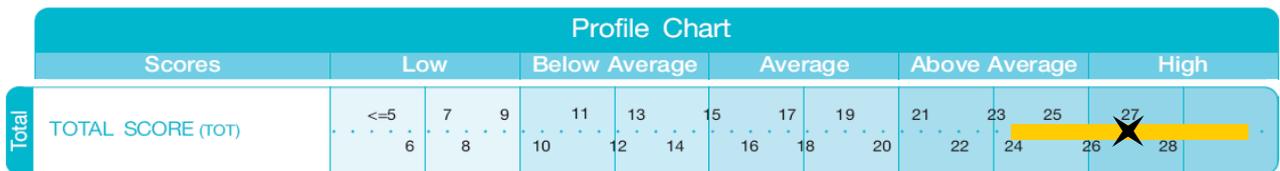
Reliability and Difference in Scores

The example below illustrates applying the principles of reliability and differences in scores in practical terms for three candidates who have completed the same test. You can see from this example that the band of error around each test score represents the Standard Error of Measurement (SEm). This is always given to you on the test Profile Chart in terms of number of T score point to draw either side of a candidate's score. In this case, the band of error is 5 T score points either side of the total score (TOT). For you to be able to say there is a difference in performance between applicants, the bands of error must not overlap.

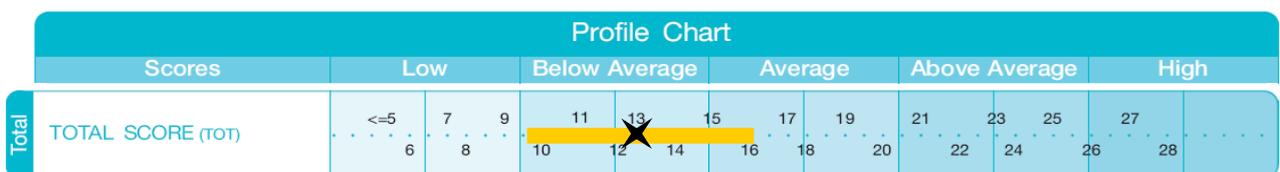
Candidate A



Candidate B



Candidate C



In the above example, three candidates have completed Professional Verbal Analysis. Candidate B has performed best, followed by Candidate A and then Candidate C. There is a significant difference between Candidate A and B, and B and C. However, there is not a significant difference between Candidate A and C in terms of their performance on this test.

Section 10: Validity

Aspects of Validity in Aptitude Tests

- **Face Validity:** do the questions 'look right' i.e. appear to be appropriate/job-relevant? (And do the reports 'look right'?)
- **Content Validity:** the extent to which the questions are actually focused on job-relevant content.
- **Consequential Validity:** the intended and unintended consequences of using a test.
- **Construct Validity:** can refer to a wide range of different sources of evidence demonstrating that a test measures an expected underlying construct, trait or theory, e.g. evidence that the test correlates with other similar measures.
 - In a sense, other forms of validity could be seen as aspects of construct validity – though often we need to look for more specific forms of validation evidence.
- **Faith Validity:** an unfounded belief that a test is appropriate, i.e. in the absence of evidence.
- **Criterion-related Validity:** evidence that the test predicts relevant criteria (e.g. competencies).

Notes:

Criterion-Related Validity

Two ways of conducting criterion-related validity studies:

- **Concurrent criterion-related validity coefficients:** the correlation of job incumbent test scores with job performance measures gathered at the same time.
 - **Pros:** quick to establish
 - **Cons:** provides a link to current performance but not directly to future performance
- **Predictive criterion-related validity coefficients:** the correlation of applicant test scores with their job performance a few months later.
 - **Pros:** provides a link to future performance
 - **Cons:** have to wait until recruits have job performance data

Problems in Criterion-Related Validation Studies

- Difficulty getting large samples
- Unreliable job performance criteria
- Motivation of participants in research settings
- Restriction of range

Notes:

Meta-Analysis and Validity Generalisation

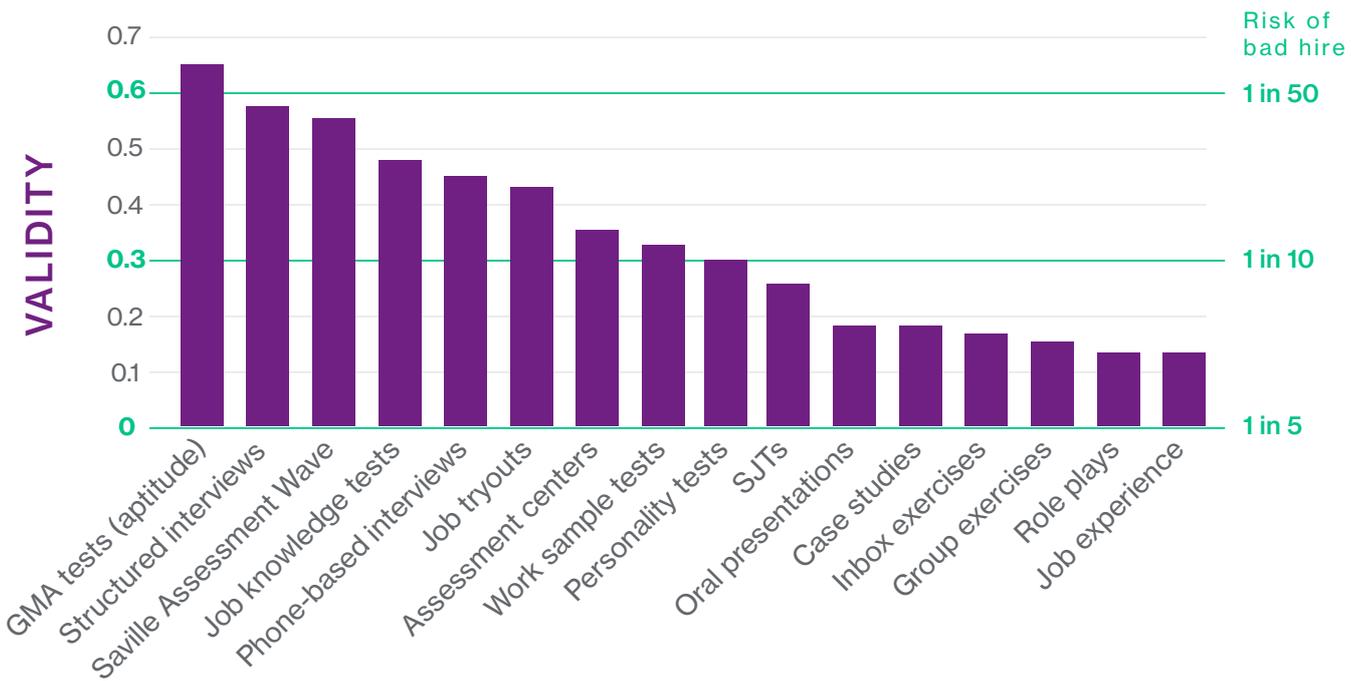
- Meta-analysis combines multiple studies to give an overall indication of the validity of different assessments
- Meta-analytic research consistently shows cognitive tests to be strong indicators of performance across different jobs
- The operational validity of ability tests is approximately +.5, and generalised across all jobs (Hunter and Schmidt 1998, Smith et al 2001) and many different groups
- This forms a very strong argument in favour of using tests

Notes:

Validity of Different Selection Techniques

Effectiveness of assessment methods*

*Includes all assessment methods generally deemed acceptable for use in hiring across different occupations

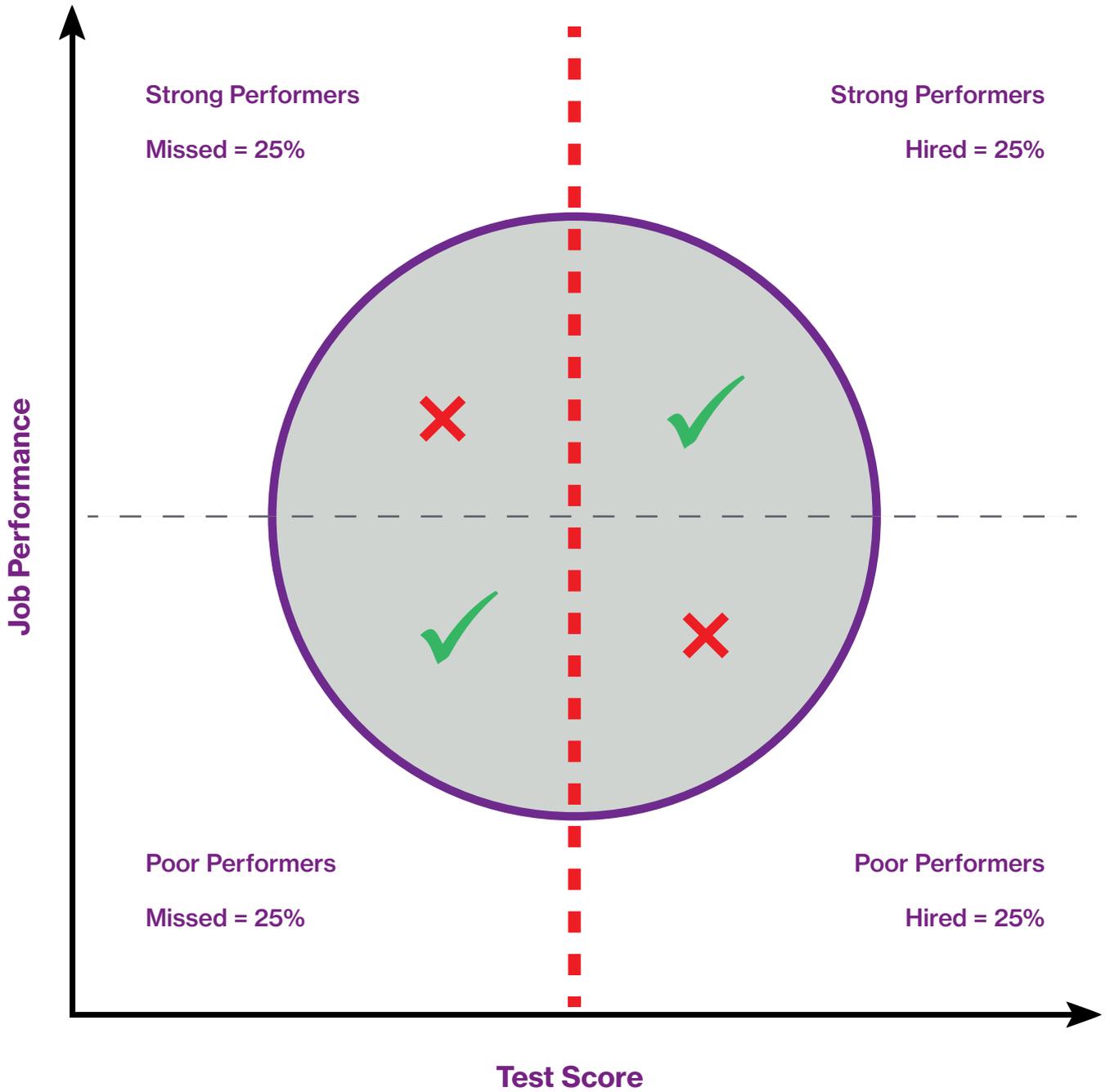


Hunter & Schmidt (1998), Schmidt et al (2016) and Saville et al (2012)

Notes:

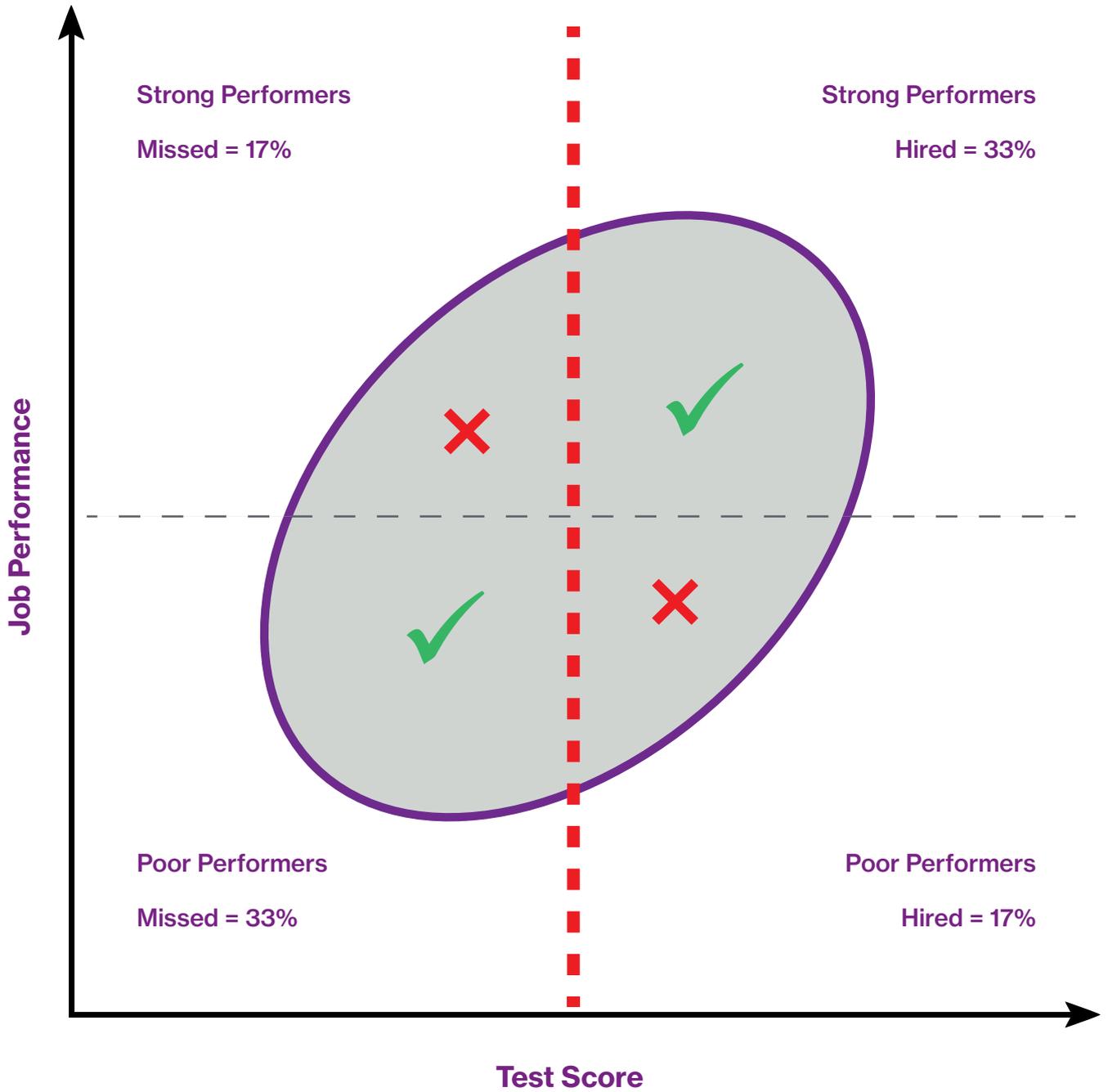
Better Validity = More Hits, Fewer Misses

$r = 0$



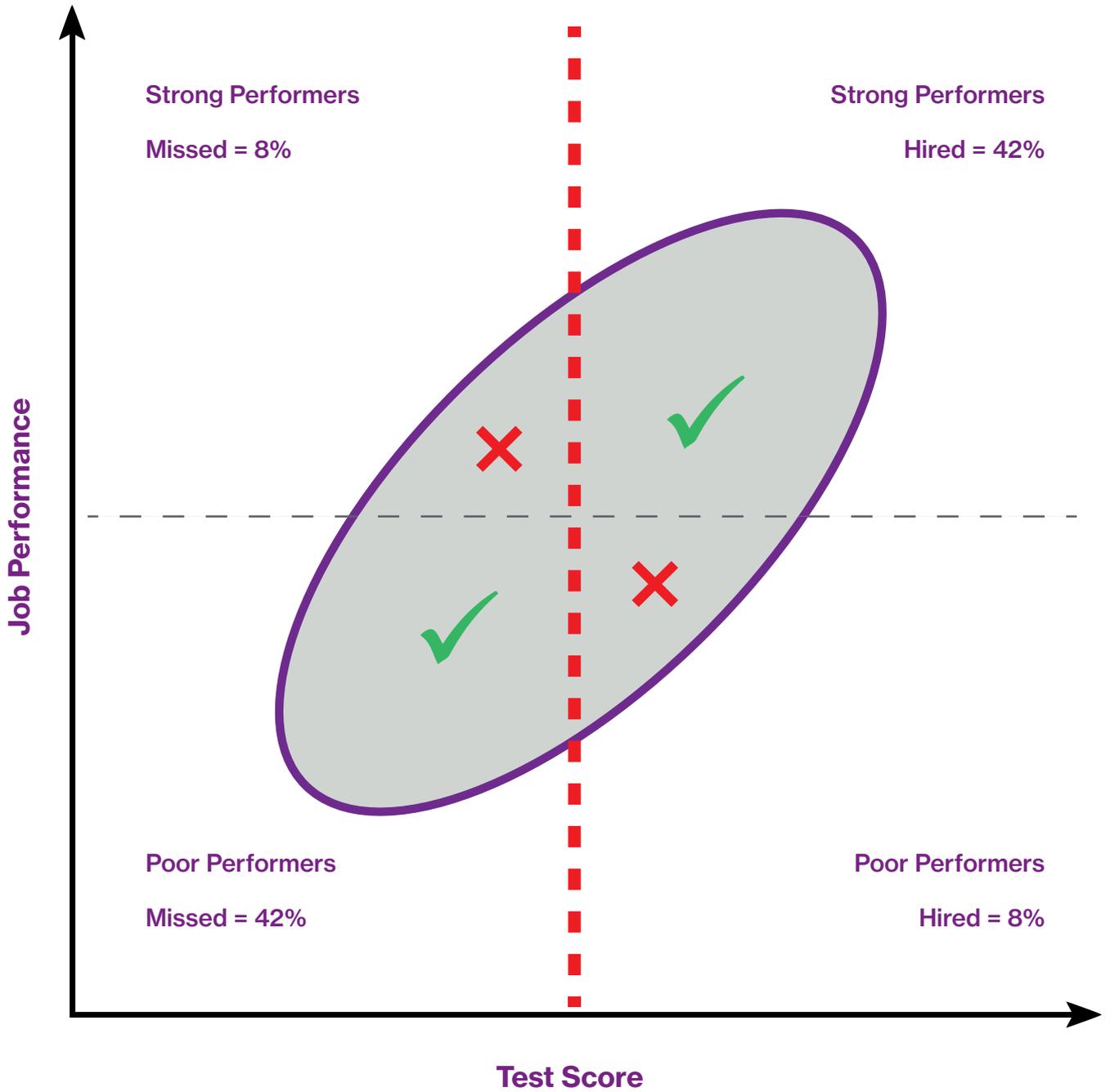
Better Validity = More Hits, Fewer Misses

$r = 0.3$



Better Validity = More Hits, Fewer Misses

$r = 0.6$

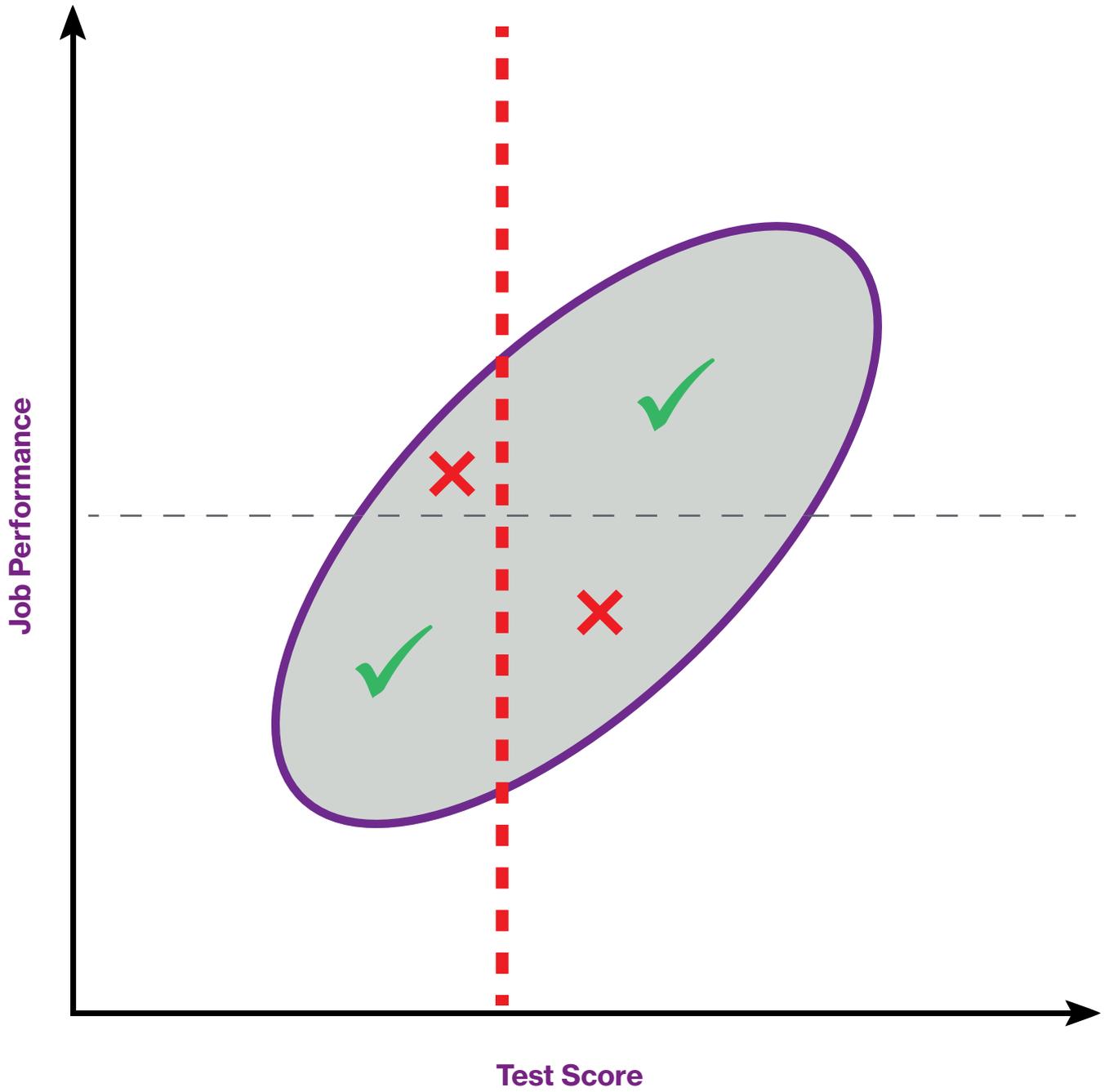


What is Test Utility?

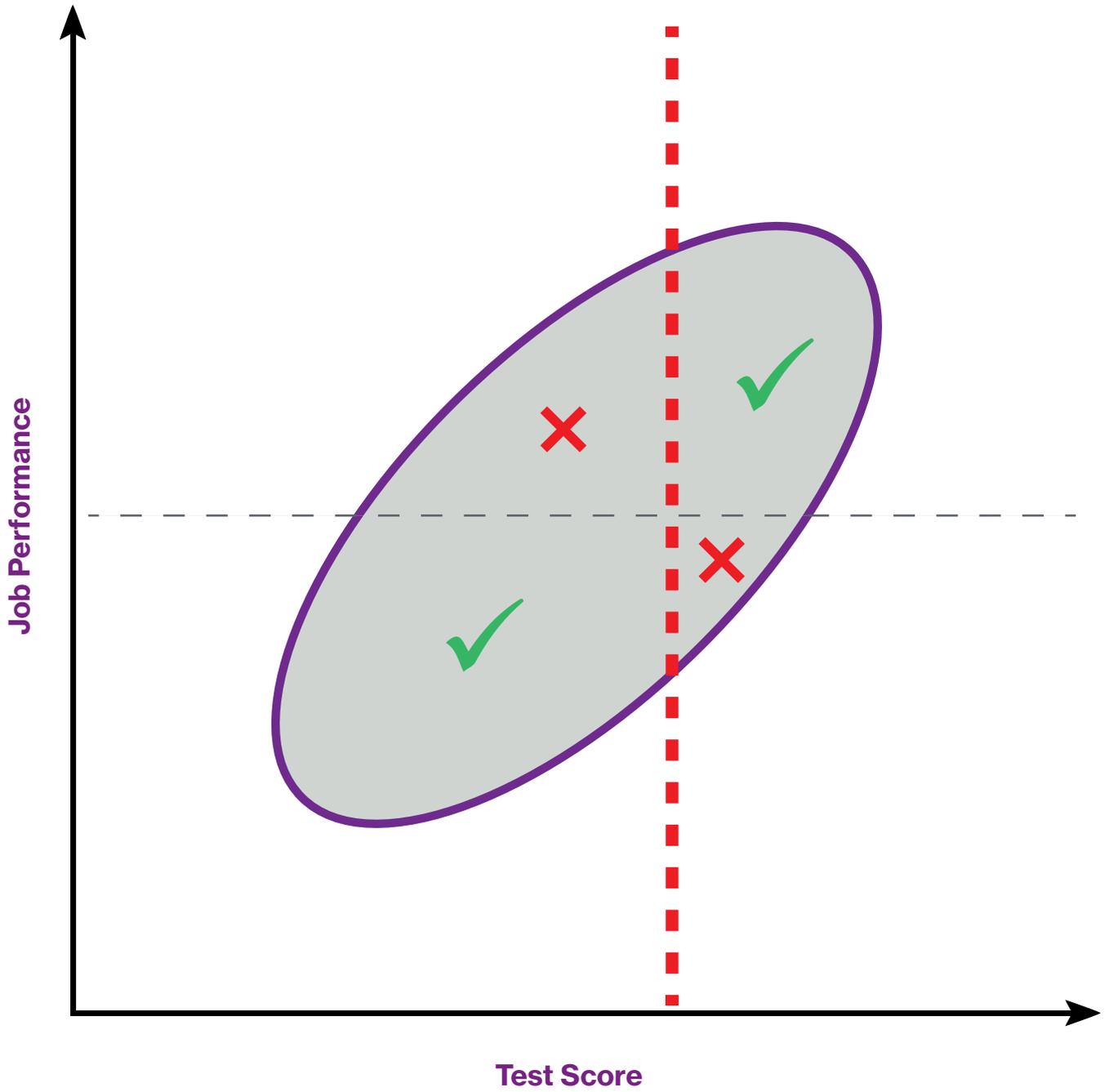
- Test utility is the return on investment from using tests that can be calculated in some circumstances – e.g. sales jobs
- Test utility is maximised when:
 - The **validity** of a test is high
 - There is high **variability** in job performance: i.e. some perform really well and some perform poorly
 - The test score can be used to **cut out high numbers** of candidates: e.g. a cut-off at the 70th percentile or above

Notes:

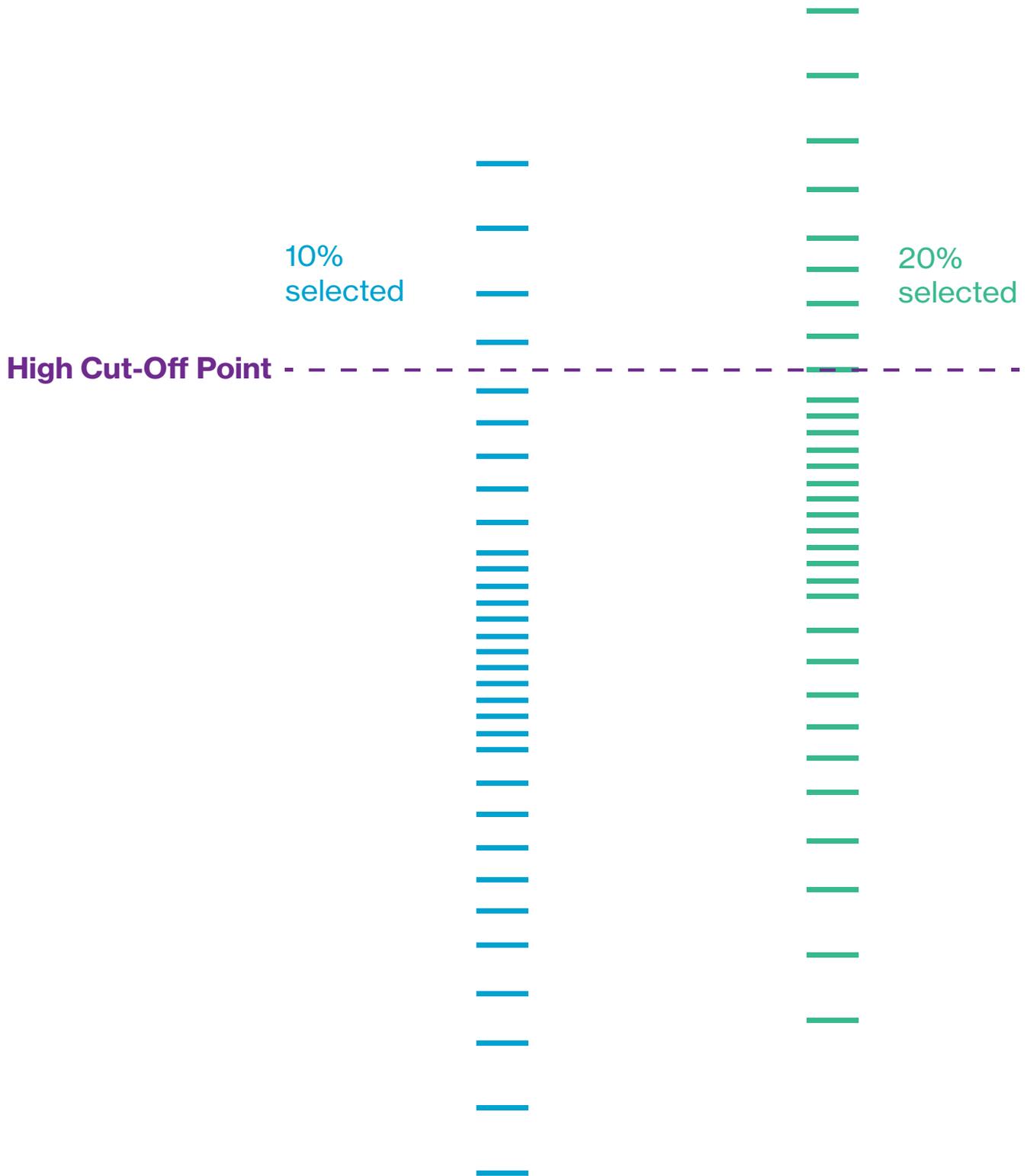
Typical Low Cut-Off e.g. 30%ile



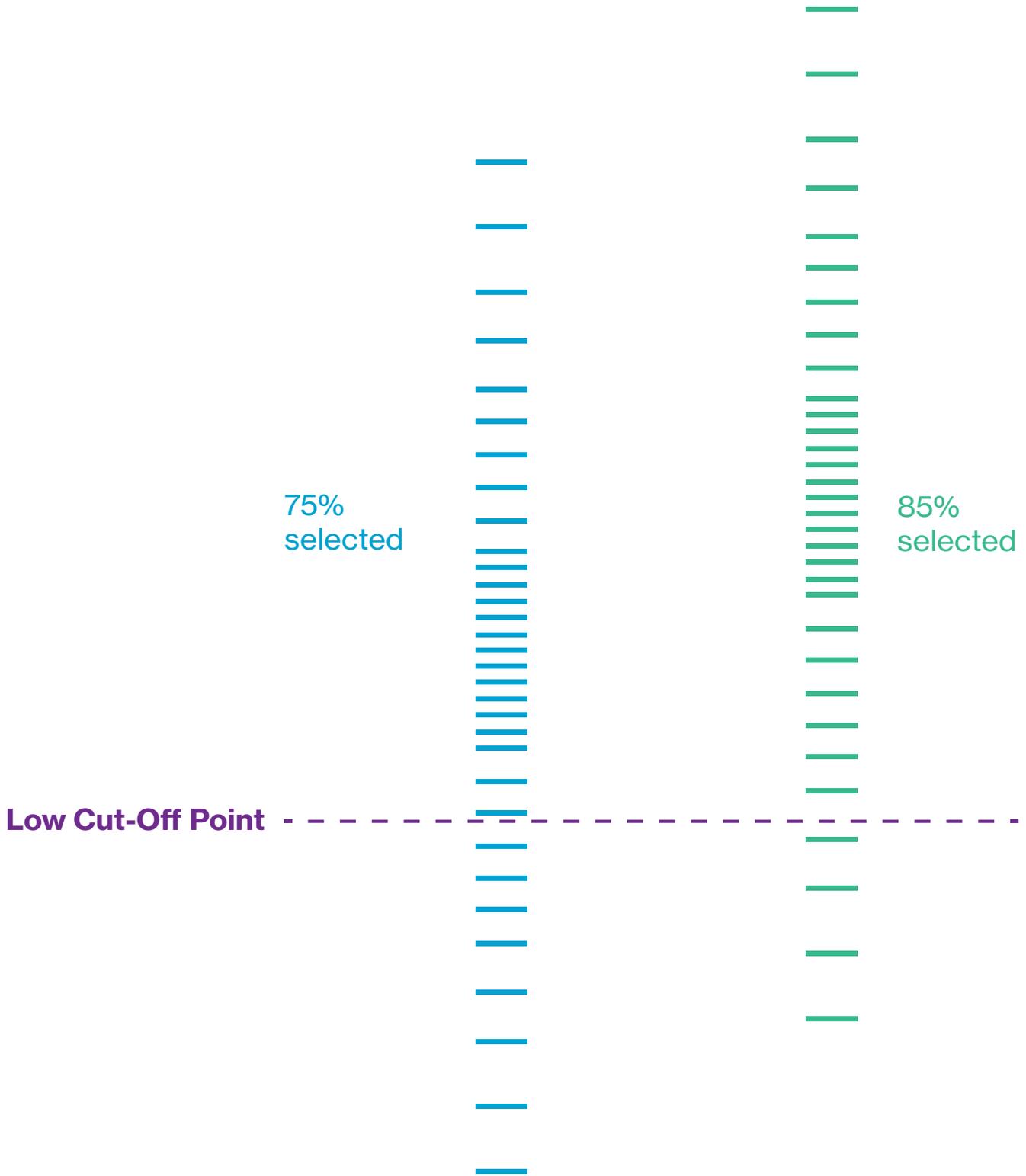
Unusually High Cut-Off e.g. 85%ile



Fairness Using Cut-Offs – With a Group Difference

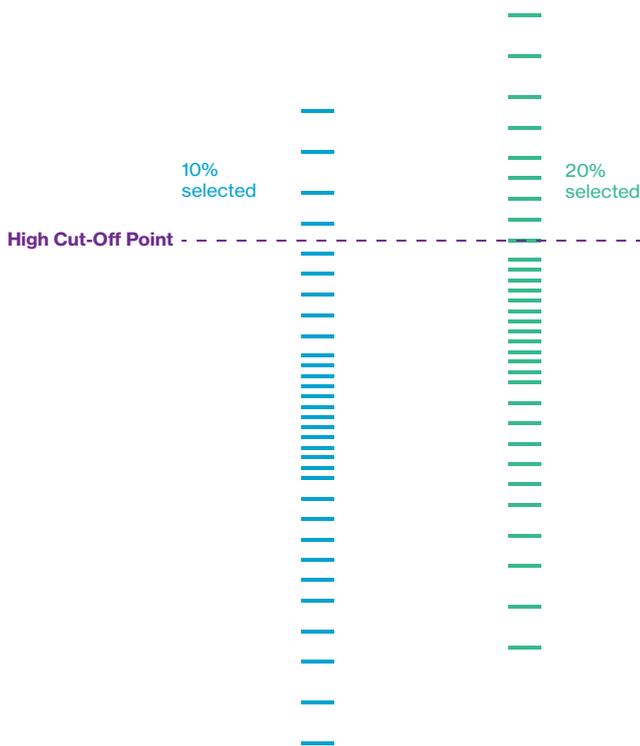


Fairness Using Cut-Offs – With a Group Difference



The Four-Fifths Rule

- A selection rate for any protected group which is less than four-fifths (or 80%) of the rate for the group with the highest rate of selection is seen as adverse impact
- This affects where we need to set cut-off scores in assessment



Blue group 50% of Green group:
Unacceptable Adverse Impact



Blue group 88% of Green group:
Acceptable Adverse Impact

Setting Cut-Offs

- Lower cut-offs are often effectively applied at early screening stages of volume recruitment (e.g. 30th%ile)
 - Still need link to job performance (job analysis and/or criterion-related validity)
- Higher cut-offs can give even greater return on investment but require more justification (i.e. evidence the job is difficult or that there is strong validity)
 - Higher cut-offs can result in lower proportions of potentially disadvantaged groups being selected
 - Particular care should be taken when using top-down selection

Notes:

Case Study: Validity

You managed to convince the hiring manager of the reliability of your Swift test, however now they are challenging how useful it is. They tell you that they get a feel for people quickly and all it takes is a cup of coffee with someone to determine whether they'd be good at the job in question. They tell you that they don't believe that tests are relevant to their recruitment.

You must present a case to the hiring manager for including your Swift test in the online unsupervised part of your assessment process.

Q1. Use the technical documentation to find validity estimates for the Swift test chosen.

Swift Analysis Aptitude or Swift Comprehension Aptitude

- Criterion-related validity:

Q2. Write some notes on how you would present your case. You should include reference to the established validity of aptitude tests and the job requirements.

Notes

What is Test Validity?

A test is valid to the extent that it measures what it is designed to measure. In particular, validity is a measure of how relevant a test is to job content. This is a key aspect of using occupational tests; if the test is not valid, then there is little point in using it. You may have a highly reliable test, but if it is not measuring the particular job competency you are interested in assessing, then it is not useful. Remember, that a valid test has to be reliable in the first place.

This section of the course will equip you with an understanding of validity and validation research. On this basis, you could potentially carry out your own validation study, but Saville Assessment offers consultancy services to conduct such research on your behalf, covering all the stages from the planning, data gathering, analysis and presentation of results, as well as recommendations for your organisation.

Types of Validity

Face validity concerns the degree to which the test looks right, i.e. whether it is relevant and acceptable. This is the reason why Saville Assessment has different test portfolios (Administrative Aptitudes, Commercial Aptitudes, etc.). The tests assess the same skills (e.g. Error Checking) at the same level but in different contexts. Tests with high face validity ensure buy-in from candidates and line managers, but with face validity test choice is not based on hard evidence and is not legally defensible if challenged. However, it may be the lack of face validity which ignites a legal challenge when candidates question the relevance of the questions they are being asked in relation to performing effectively on the job.

There is a danger that users may rely on spurious validity, such as face validity, as evidence of its true validity. It cannot be assumed, for example, that because a test is face valid, that it is also psychometrically valid. Using a test that is not psychometrically robust can subsequently lead to mistakes in selection, development, feedback and interpretation.

Content validity is a more rational basis of choosing tests. Test choice is preceded by job analysis/role profiling activities and tests are chosen based on the degree to which they assess the key skills required for the job. The closer the test replicates the tasks required on the job,

the better. If you have established content validity, then by definition you have face validity. Content validity is legally defensible.

Consequential validity considers the intended and unintended consequences of using a test. For example, if an assessment is being used to identify high potential people within an organisation for succession planning purposes, intended consequences could include encouraging individuals to strive to develop themselves in performance-relevant areas, greater motivation and effort displayed by potential succession candidates and improved understanding of what matters for effective performance. On the other hand, unintended consequences could include a narrowing of focus amongst potential succession candidates to just those variables assessed by the assessment, potential succession candidates engaging in practices to disadvantage others and inappropriate use of assessment scores by the administrators or decision-makers.

Construct validity concerns the extent to which a test measures some underlying theoretical construct, e.g. does the test really measure mechanical comprehension? Construct validity can be established by correlating a new test with existing tests designed to measure the same theoretical construct.

Faith validity concerns the degree to which a test user has a feeling that the test works – this is similar to interviews, where interviewers state: “I can spot a good candidate the minute they walk in the door.” Test users can have an unquestioning belief that the test works, and this can be the basis of some organisations using the same tests for long periods when there may be more appropriate and up-to-date tests. Faith validity can aid in getting buy-in to the use of objective assessment methods, but as decisions regarding test usage are not based on any hard evidence, it can lead to misuse of tests and in the worst-case scenario could lead to the use of tests that are not legally defensible or valid (so not allowing for the selection of better candidates).

As practitioners, we are more interested in the practical relationship between test scores and job performance. **Criterion-related validity** is the most powerful form of validity, as it focuses on correlating test scores and job performance measures to identify the degree to which test scores predict performance. There are two main types of

empirical validity, which are similar in approach but different in terms of the timing of data collection.

- **Concurrent validation** studies involve the collection of test scores and job performance measures at the same time, typically with jobholders.
- **Predictive validation** studies involve collecting test scores from job applicants, waiting for a period of time while work experience is gained, and then the collection of job performance measures.

The benefit of carrying out criterion-related validation studies is that you get very powerful information regarding how well tests are working and they provide sophisticated information on cut-offs and weightings. This approach is also legally defensible in terms of justifying test choice, if challenged.

Key Problems in Validation

Carrying out validation studies is not without its drawbacks. To ensure that you have a representative sample, it is recommended that you find at least **100** jobholders for a concurrent validity study or **100** applicants for a predictive validation study. That said, 200 is better - the more the merrier (or the more confident we can be in interpreting the results meaningfully).

The criterion problem: One of the key challenges when carrying out validation studies is being able to measure job performance effectively. Unfortunately, measuring job performance is very difficult and this is known as the criterion problem. It is often impossible to get one good measure that encompasses how effective performance in a job as a whole is. It is preferable to identify objective, quantifiable measures, such as quantity of output on a job, promotion rates, financial results, e.g. sales performance, work sample tests or training course results. Alternatively, the use of self and boss ratings can be considered. Many assessments of job performance such as Saville Assessment's Performance 360 now measure a number of different dimensions of performance at work, based on the boss, colleague and report ratings. Saville Assessment have a range of performance rating questionnaires that can be used by organisations wishing to carry out validation studies.

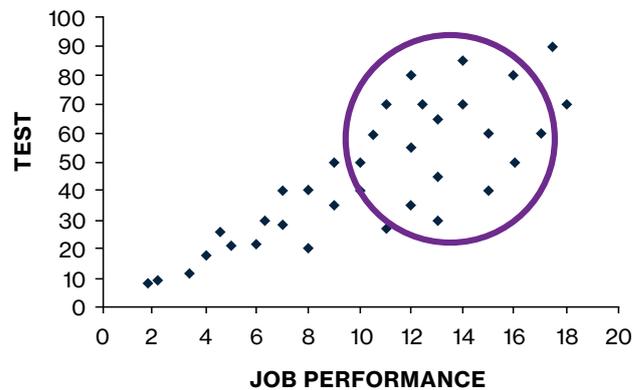
The problem of collecting objective and reliable measures of job performance presents a real challenge when carrying out validation studies. Regardless of the job performance criteria which are chosen, there will always be error that leads to a reduction of the correlation of tests/ other techniques with the criteria. There is a formula to

correct for unreliability of the criterion data – this is called correction for attenuation.

Restriction of range: Ensuring that the sample of people taking part in the study is motivated and representative of the population of applicants or jobholders can also be difficult. Samples of jobholders taking part in concurrent studies should ideally comprise a range of performers, including high, moderate and less effective performers. However, because jobholders have gone through a selection process, the range of performance may be limited and observed correlations between test scores and job performance can, therefore, be reduced. This is termed restriction of range.

The graph above shows a scattergram, cut from an example validation report. It shows test scores (from 0 = 100) on the vertical x axis and a measure of job performance on the horizontal y axis (based on a rating scale from 0 – 20).

In fact, this is an unusual situation, as what we are looking at in the full graph is where a test has been given but not used to make a selection decision.



The circle indicates what would happen if the test was used and where the lower scorers on the test were not selected and perhaps the person who was high on the test chose not to take the job due to accepting a better offer. This gives a picture of what can happen when the range of scores is used in a practical situation to make selection situations (in this case we would not have the data on the people outside the circle, we would only see those within).

The circle magnifies a section of the scattergram which represents a smaller subgroup of individuals within the total group.

If you were to carry out a correlation calculation using just the scores in the circle, the resulting validity coefficient

would actually produce a lower correlation than if you had based your calculation on the whole sample group. Thus, by restricting the range of scores artificially, by only taking top scorers and validating on them, you will produce an underestimate of a test's validity.

Restriction of range can be corrected for using a formula.

Meta-Analysis and Validity Generalisation

Validity generalisation is the extrapolation of validity correlations established in research to other settings. When seeking to generalise the findings from one validity study to another situation, this should be done on the basis that the situation is similar to that of the study e.g. if the study demonstrated a numerical test was predictive of the performance of sales representatives, we would be more confident in applying this test to the recruitment of sales advisors. Alternatively, if the study is linked to similar criteria to what we are trying to predict, we would have more confidence in using the test e.g. commercial problem solving.

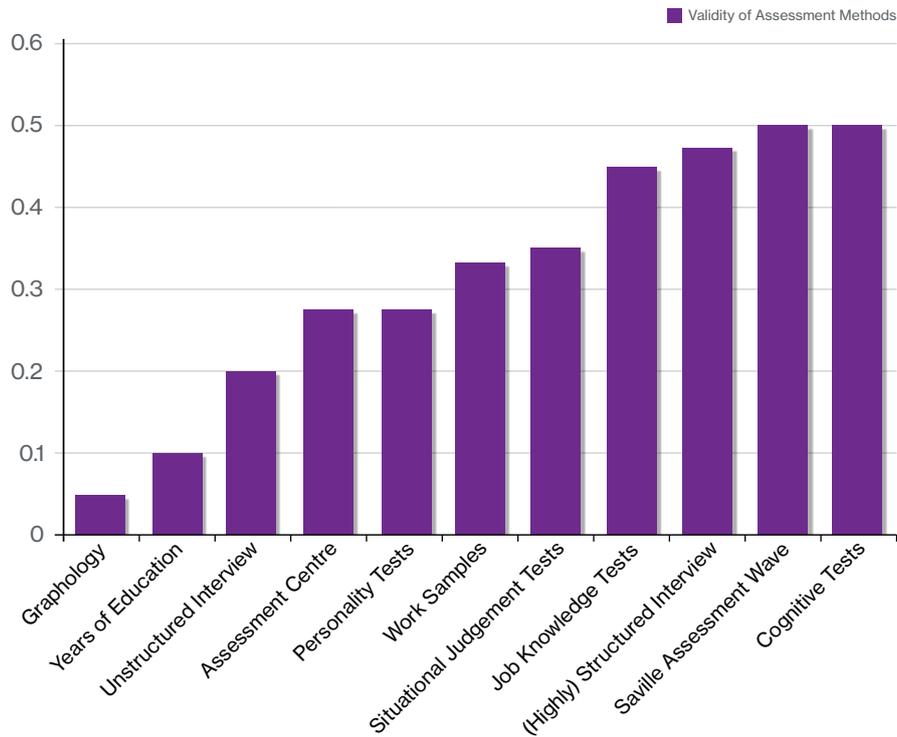
Meta-analysis has revolutionised our capability to make well-reasoned decisions based on a large number of research studies which often seem contradictory (i.e. give seemingly different results and come to seemingly different conclusions). Meta-analysis is a statistical process which combines many studies which have related research aims so as to form overarching conclusions; by increasing the sample size you minimise the chances of sampling error.

Using meta-analysis, **Hunter and Schmidt (1998)** combined many studies looking at how well a variety of different selection techniques predicted performance. This process enabled them to correct for sample size, unreliable job performance criteria and restricted range. They found the validity of ability tests is approximately +.5 and generalised across all jobs.

Salgado and Anderson (2002) on a sample of 75,000 people in Europe also found that ability tests generalised across all jobs with no evidence of differential validity. Tests of specific abilities were just as valid as tests of 'g'. The more complex the job, the higher the correlation of ability tests with job performance.

Notes:

What Works?



The validity scale on the graph above ranges from 0 to 1, where 1 is perfect, i.e. if a selection method scored a validity of 1, it would be 100% perfect in predicting job performance. As a guide, a validity score of 0.4 would be a 'strong' score on these methodologies which are quite difficult to correlate with job performance. As shown above, references and educational qualifications were not very predictive. Structured interviews were more predictive than standard interviews (0.45 vs 0.2). Specific cognitive ability tests were the single highest predictor at 0.51.

Test Utility

Cost-benefit or test utility is concerned with answering the question: what is the pay-off from using tests? Ultimately this can be viewed as the most convincing argument for using tests. Utility equations take account of the costs of testing versus the potential gains of using tests. The more valid the test and the higher the cut-off on the test, the greater the utility. In addition, higher the variability in job performance (i.e. some perform really well and some perform poorly) and the more people that can be cut out using the test, the greater the gain. The benefit is linear to the validity coefficient, with more valid tests giving greater benefit. You should, therefore, choose more valid assessments over less valid assessments.

Implications for Utility

There are implications for the selection strategy used by organisations, based on the concepts introduced by Taylor and Russell back in 1939 in what is still one of the most famous papers in industrial and organisational psychology.

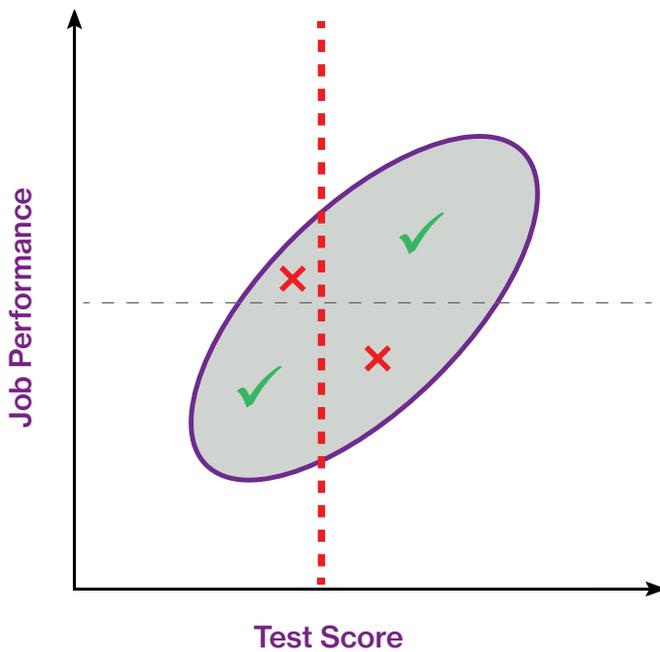
Using low cut-offs:

In situations where organisations can't be overly selective, perhaps because there are roughly the same number of jobs per applicant or sometimes even less applicants than jobs, using low 'cut-offs' is likely to be the most appropriate strategy. In this instance, a relatively low percentile (e.g. 30th percentile) would be chosen as the point at which an applicant needs to score above in order to progress. Those who are not up to this standard are screened out. The graph below demonstrates how, in these instances, whilst you are likely to reject fewer people who will be successful, you are also more likely to hire more people who will fail. This is likely to lead to lower overall test utility/cost benefit as the overall calibre of the workforce is likely to be lower. To combat this, hirers should seek to improve their ratio of candidates to roles available by running effective attraction campaigns, which will give them more choice about who they hire in subsequent rounds of assessment.

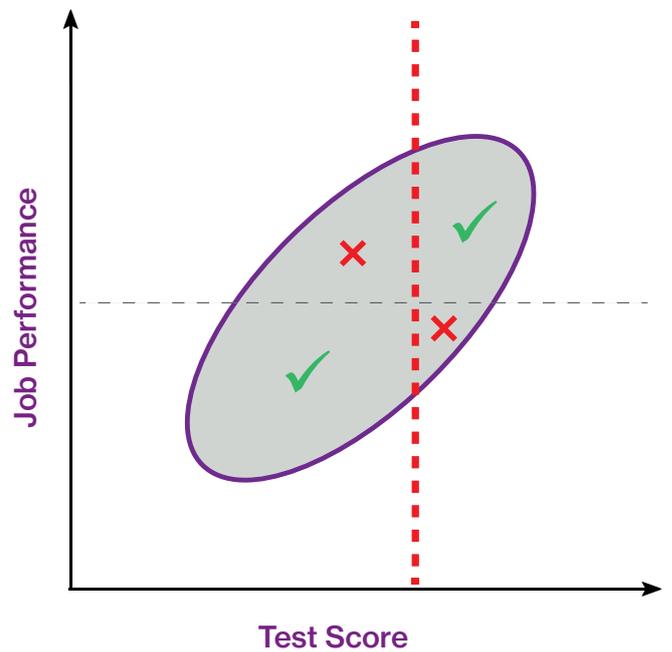
Using high cut-offs:

Where organisations can be highly selective because there are many applicants to few available jobs, the overall utility or cost benefit is likely to be high. In this instance, the number of people who get hired but are likely to fail significantly reduces. The price you pay for this, however, is rejecting people who would have succeeded if they were hired. And while this usually presents more of a problem for applicants than hiring managers and organisations, as is becomes more and more important to get the candidate experience right, thought should be given to how to deal with this group.

Typical Low Cut-Off



Unusually High Cut-Off



The sort of approach that might be put in place where you have a lower selection ratio (and so a higher number of applicants than jobs available), is a ‘top-down’ strategy whereby top scorers on the test are selected sequentially, until all the jobs are filled. This raises the calibre of the workforce as only top scorers are selected. However, there are implications for equal opportunities in this situation, as applicants from lower scoring groups are less likely to be selected.

To illustrate this point, consider the two diagrams on the next page. In both cases, there are two groups of people who have completed the same assessment. Each blue line represents the score of one person in the “Blue Group”, with higher lines representing higher scorers. Likewise, each green line represents the score of one person in the “Green Group”, again with higher lines representing higher scorers. As can be seen in both diagrams, the Green Group is, on average, of a higher calibre than the Blue Group, although there is considerable overlap in performance across the two groups (that is, there are many people in the Blue Group who have outperformed people in the Green Group and vice versa). Both groups also show a bunching of scores in the centre of their range around the average score, with fewer very low and very high scores. This is what we’d expect to see in data sets where assessment scores are normally distributed.

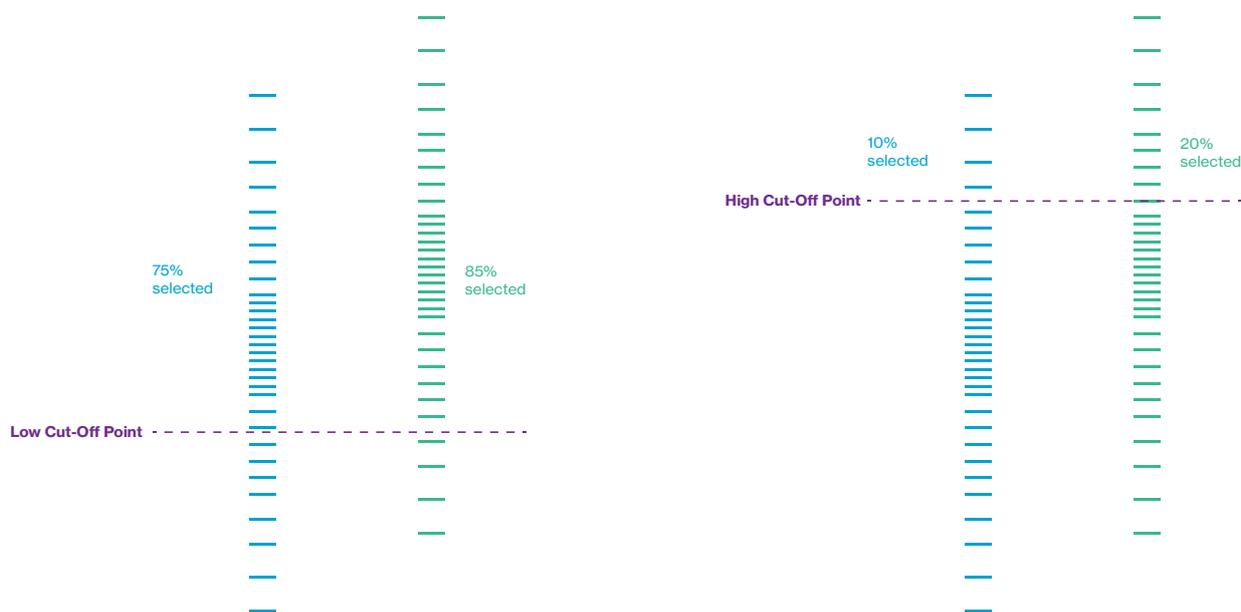
In the first diagram, a low cut-off score has been used, as is represented by the dotted line. Anyone who scores below this line is cut from the assessment process (i.e. they haven't achieved the required score to be considered further). The implication of this particular cut-off score is that 75% (or 0.75) of the Blue Group have achieved the required cut score, whereas 85% (0.85) of the Green Group have achieved the required cut score.

However, in the second diagram, a high cut-off has been used. Again, this is represented by the dotted line. In this second case, only 10% (0.10) of the Blue Group have achieved the required cut score, whereas 20% (0.20) of the Green Group have achieved the required cut score.

A particularly important point about setting high cut-off scores is that they can increase the chances of disproportionately cutting out candidates from one group versus candidates from another. There can be a risk of falling foul of equal opportunities legislation in such cases. One way in which the possibility of one group being disproportionately impacted by cut-off scores (sometimes known as adverse impact) is the 4/5th or 80:20 rule. Here, the relative rate of different groups achieving the cut-off score are compared. Where the pass rate for one group is seen to be less than 80% (or 4/5th) the pass rate of another group, this may be seen as evidence of adverse impact, which could indicate unfairness in the assessment process.

In the examples below, we can see that the lower cut-score wouldn't break the 4/5th rule (because 0.75 is more than 4/5th of 0.85), whereas the higher cut-score would break the 4/5th rule (because 0.10 is less than 4/5th of 0.20).

In summary, reducing the level of cut-offs can reduce the risk of adverse impact or unfairness in an assessment process. Nevertheless, it's still important to bear in mind that every sample of people is different and any apparent evidence of adverse impact or unfairness may be because of sample-specific factors. Thus, such evidence shouldn't be excessively generalised. Similarly, if there appears to be adverse impact in a data set, reducing the cut-off set is likely to be a more effective solution than throwing the baby out with the bath water and changing the assessment itself.



Section 11: Approaches to Online Testing

Security of Test Content

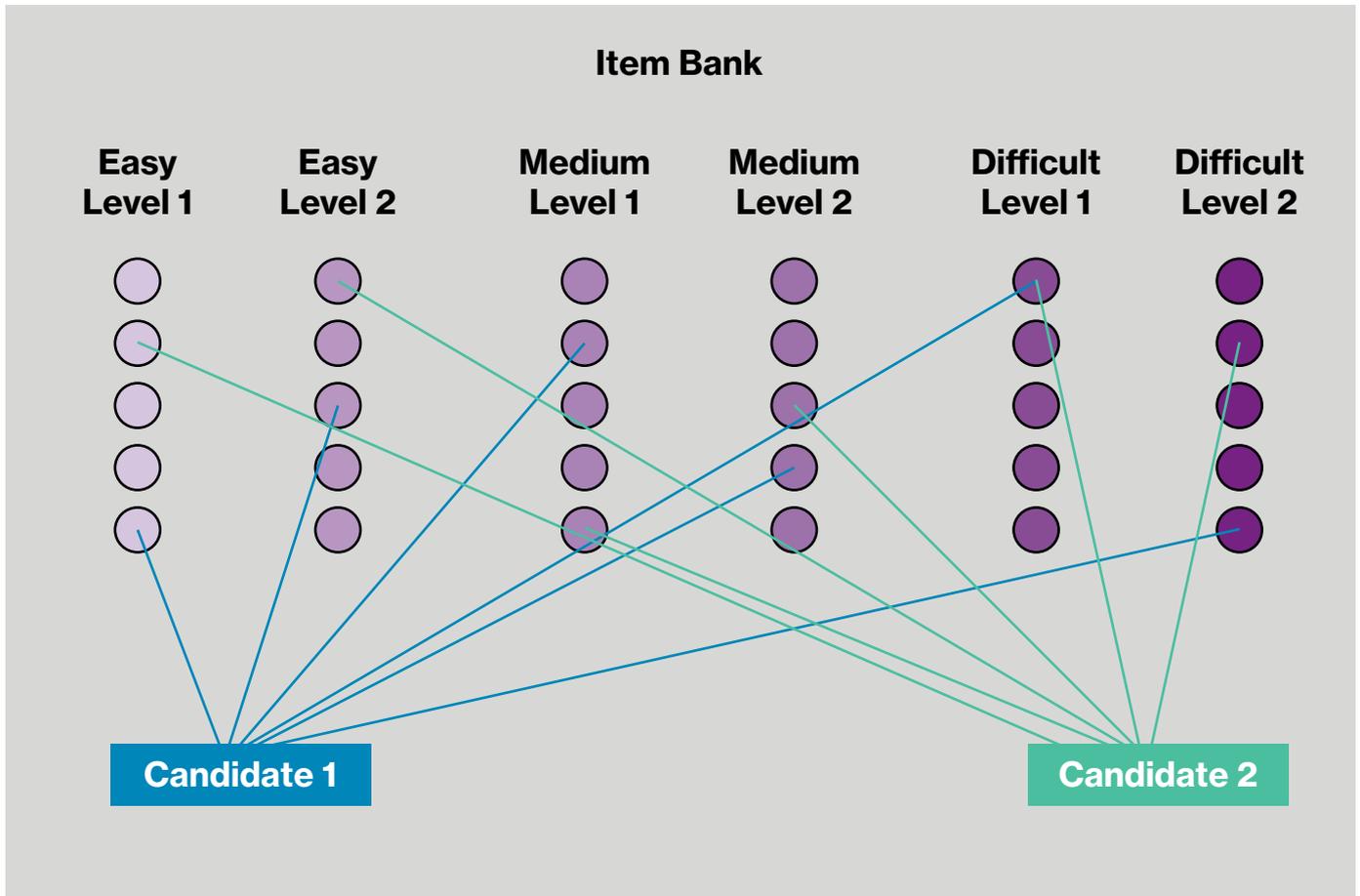
- With all testing there are potential issues relating to the security of test content
 - This is particularly true for remote online testing
- One method for reducing security issues is to present different content to different candidates by holding content in a bank
- Item Response Theory (IRT) provides a method to score candidates in a consistent way, despite them seeing different test content
- Classical Test Theory (CTT) is more generally used for fixed-content testing
 - In CTT the number of questions answered correctly = ability
 - However, this assumes all items are of equal difficulty

Item Response Theory

- IRT provides a method which takes account of the difficulty of each item within a test
- The **Test Information Function** allows us to see what level of candidate ability (low, average or high performance) is most accurately measured by a particular test
- Two common approaches using item banking are **Gradient Step Testing (Fixed Length)** and **Adaptive Testing**

Notes:

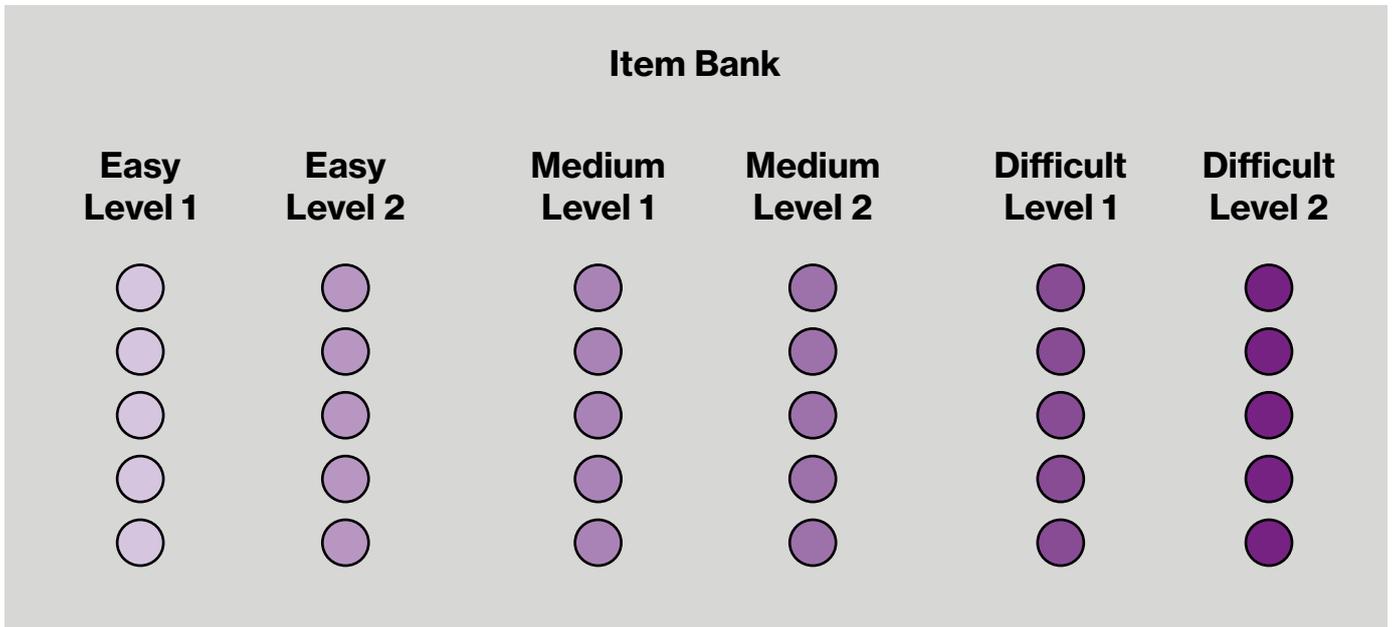
Gradient Step Testing (Fixed Length)



Items are randomly drawn from the bank and the number of items and difficulty level remains consistent. Each candidate would do a six-item test of equal difficulty.

Notes:

Adaptive Testing (Uncontrolled Length)



Candidate 1

Easy Level 1	○	✓
Easy Level 2	○	✓
Medium Level 1	○	✓
Medium Level 2	○	✓
Difficult Level 1	○	✗
Medium Level 2	○	✗
Medium Level 1	○	✓
Medium Level 2	○	✓
Difficult Level 1	○	✗
Medium Level 2	○	✓

Candidate 2

Easy Level 1	○	✓
Easy Level 2	○	✗
Easy Level 1	○	✓
Easy Level 2	○	✗

With Adaptive Testing, items are randomly drawn from the bank as the test narrows in on the ability level of the candidate. The number of items in a test will be different for different candidates.

Why Use Gradient Step Testing (Fixed Length)?

Well designed GST/fixed length tests are short, fair, valid, reliable and ensure a positive and consistent testing experience.

- Candidates are given tests of the same difficulty
 - Therefore, there is a greater perception of consistency
- Candidates have the same time to complete the test
 - Therefore, they feel they have an equal opportunity to show what they can do

Notes:

Notes

This section gives you an outline of what is involved in developing a new online test. It involves consideration of **Item Analysis, Item Response and Item Banking**. This section can also help you to evaluate tests that you are considering using, enabling you to understand whether they have been constructed correctly. The section also gives you an indication of the complexity of issues involved in designing a good psychometric test.

New tests are being designed to reflect the growing need to assess high volumes of applicants in an environment where candidates are completing them in unsupervised settings. As such, the security of test content is now often at the forefront of test construction.

Item Response Theory (IRT) is increasingly used in the design of ability tests, although its origins are pre-1950. **Classical Test Theory** has very few assumptions and is very robust i.e. a person's obtained raw score = true score + error. IRT goes beyond this but has a complex theoretical basis with many assumptions that are often not met in practice (e.g. unidimensionality of item pool, unspeeded assessments). IRT equations model the theoretical ability or 'latent trait' measured by the test. Item characteristic curves show how items measure the underlying ability the test is designed to measure – like item scale corrections in classical item analysis. IRT can be used, as can Classical Test Theory, to produce many parallel versions to overcome exposure issues, but cannot ensure exactly who is taking the test. IRT tests can have considerable overlap in content as the 'best' items are picked by the algorithms repeatedly. IRT can be used to produce 'Adaptive' tests which jump to more difficult questions if the candidate answers easier items correctly and to easier items if the candidate gets the answer wrong. However, adaptive tests may not always save on testing time and may be less popular with candidates as they will often be presented with highly challenging items very early on within the test. There are also legal concerns about the fairness of the process if candidates are asked different questions to one another. There is no evidence that IRT produces more valid tests than well-written tests using classical test theory. Nevertheless, IRT sensibly used has a part to play, but is not the panacea sometimes proposed – security remains a very real problem.

Item Response Theory

Tests are commonly scored using classical test theory where, normally, the number of correct answers are added together to create the total score.

In classical test theory, the formula ' $X = T + E$ ' is used to describe what is being measured. X is the observed score, i.e. the score the individual has on the test, T is their true score and E is the error. In IRT, ability estimates are calculated rather than observed scores. These ability estimates look at the person's underlying ability which is regarded as independent of the test.

For CTT the sample used to calculate item statistics needs to be representative of the intended test users; this is not necessary for IRT.

With IRT, for example, item statistics are calculated for three different variables: discrimination (the information provided by the item), difficulty and pseudo-guessing. These parameters are used to calculate the probability of an individual getting an item right or wrong for a given ability, for example, if an individual answers a difficult item correctly their ability is likely to be very high; if they get an easy item correct their ability is unlikely to be very low, but it could range from fairly low, to moderate or even very high.

There are three IRT models, the 1PL model (or 'Rasch' model named after its creator), the 2PL model and the 3PL model. The 1PL model only looks at difficulty and assumes all items provide equal amounts of information. The 2PL looks at difficulty and discrimination and the 3PL looks at difficulty, discrimination and pseudo-guessing. This can be used for modelling multiple choice items where there is a fair chance of guessing the correct answer.

The Test Information Function is an extremely useful feature of IRT. It basically tells you how well the test is doing in estimating ability over the whole range of ability scores. A test is a set of items; therefore, the test information at a given ability level is simply the sum of the item information at that level. The general level of the test information function will be much higher than that for a single item information function. Therefore, longer tests will measure an examinee's ability with greater precision than will shorter tests.

Varying Item Content from a Bank

For Gradient Step Testing (Fixed Length) tests there is randomisation in the selection of the items sat. The item selection is pseudo-random rather than truly random as there are likely to be constraints on the items selected to maintain fairness (ensuring that there is a spread in the difficulty of items or the length of time that they take).

Computer Adaptive Testing is a complex method which requires very large item banks. Candidates sit items randomly selected from an item pool that are usually of medium difficulty (although this can be adjusted where information about the likely ability of the group is available prior to testing), and then further items are selected based on their responses. As their ability is re-estimated after each item is presented only items that will provide further information about the candidate are selected. This means that test length can be reduced by up to 50%.

Notes:

Section 12: Best Practice & Ethics

Best Practice: Key Points

- Promote proper data management
- Ensure fair assessment
 - Test validity
 - Preparation materials
 - Monitoring differences
- Consider the candidate's needs
- Make accommodations for special requirements
- Review your testing policy

Data Implications for Testing

General Data Protection Regulation (GDPR)

Six principles to ensure that information is handled properly:

1. Processed lawfully, fairly and in a transparent manner
2. Collected for specified, explicit and legitimate purposes and not further processed for an incompatible purpose
3. Adequate, relevant and limited to what is necessary in relation to the purpose
4. Accurate and, where necessary, kept up-to-date
5. Kept in a form which permits identification of data subjects for no longer than is necessary for the purpose
6. Processed in a manner that ensures appropriate security of the personal data

Notes:

Equality Act 2010 (UK)

Nine Protected Characteristics

- Age
- Disability
- Sex
- Race
- Religion or belief
- Pregnancy & maternity
- Sexual orientation
- Gender reassignment
- Marriage & civil partnership

Avoiding Discrimination in Testing

- **Direct discrimination** is deliberately treating people differently because of the group they belong to
 - e.g. rejecting a candidate on the basis of their age
 - Generally not related to aptitude testing unless tests are used inconsistently (e.g. different norm groups/cut scores)
- **Indirect discrimination** is where an unjustifiable requirement or condition has a disproportionate impact on one or more protected groups
 - e.g. unjustifiably screening out more female than male candidates through the use of a particularly high cut-off on a mechanical reasoning test for an office administrator role in an engineering company

Notes:

Group Differences on Cognitive Tests

- Some differences are consistently seen between groups; age, gender and ethnicity, for example.
- Test bias can occur with items which are oriented towards one particular group's prior knowledge (this should be avoided).
- Some observed differences are not attributable to group membership or genetics but reflect environmental differences in other variables, e.g. access to good education.
 - Some differences are likely to be stable and long term, others may be temporary
 - Some differences may be interactions where the environment can impact to maximise or minimise a particular genetic predisposition

Adverse Impact in Testing

- Adverse impact is where a substantially different proportion of one group is hired in comparison to another group.
 - This is likely to occur where there is a difference between groups in terms of their average test score.
- As group differences exist, justification for using tests and the level of cut-off matters.
 - A high cut-off of, for example, the 90th percentile is likely to reduce the relative percentage of black candidates compared to white or Chinese candidates and it is likely to require strong justification to show that this level of aptitude is required.

Notes:

Best Practice: Consider the Candidate

Notes:

- Explain the assessment process
 - Important for procedural justice
- Provide practice/preparation material
- Establish informed consent
- Establish feedback process
- Promote consistency and fairness for all candidates throughout the process
- Test in the most appropriate language (normally first language)
- Enquire about special requirements/accommodations (e.g. disabilities)

Testing Candidates with Special Requirements

- You are responsible for deciding upon an appropriate accommodation/adjustment and will need to:
 - Understand the disability
 - Seek test publisher advice as to what different accommodation options are available
 - Seek expert advice on a candidate's disability where appropriate
 - Decide how to accommodate, test as standard, or not test at all
- Typical accommodations include: time adjustment, screen reader technology, large screen/print version, braille, response assistance (sighted reader or person to record responses)
- Conversations may be required with the disabled candidate, the assessment provider and even a suitably qualified professional for that disability
- Volume testers will generally have policies in place regarding common accommodations, e.g. particular time adjustments for dyslexia upon presentation of an appropriate certificate
- A larger band of error should be applied to adjusted test scores

Notes

General Data Protection Regulation (GDPR)

The test user has a responsibility to maintain and use information collected about people in an appropriate manner. There are six principles put in place by the GDPR to make sure that information is handled properly. These state that data must be:

- 1. Processed lawfully, fairly and in a transparent manner:** The scores should be used to make fair decisions about people. This requires the use of well-chosen tests with appropriate interpretation. Ensure that candidates are provided with sufficient information about the assessment process.
- 2. Collected for specified, explicit and legitimate purposes and not further processed for an incompatible purpose:** Ensure scores are only used for the purposes for which they were collected. To use them for other purposes requires gaining further permission from the candidate. If an assessment is completed as part of a development process it is unlikely it would be appropriate to use the results for selection or promotion decisions at another time.
- 3. Adequate, relevant and limited to what is necessary in relation to the purpose:** Ensure only appropriate tests are used. Tests are not used unless the information is needed for a proper business purpose, e.g. making effective selection decisions, developing staff.
- 4. Accurate and, where necessary, kept up-to-date:** Ensure that care is taken in collecting and processing data to ensure it is accurate.
- 5. Kept in a form which permits identification of data subjects for no longer than is necessary for the purpose:** That there is a policy of deleting data once it is no longer useful. Typically test scores remain relevant for 12-24 months. After this they should be erased.
- 6. Processed in a manner that ensures appropriate security of the personal data:** Appropriate security should be in place when storing data. Appropriate technical or organisational measures should be in place to protect against unauthorised or unlawful processing and against accidental loss, destruction or damage.

Each organisation should take their own legal advice with regard to their human resource activities. Saville Assessment is not in a position to advise on legal matters.

Equal Opportunities Legislation

Equal opportunities legislation exists in many jurisdictions and, in general, care should be taken to avoid unfairly discriminating against any group.

Equal opportunities legislation has developed over time to protect more groups, with major legislative developments in the latter half of the 20th Century. This legislation has continued to strengthen and evolve to cover more protected groups.

For example, the UK Equality Act 2010 protects the following characteristics:

- age
- disability
- gender reassignment
- marriage and civil partnership
- pregnancy and maternity
- race
- religion or belief
- sex
- sexual orientation

Discrimination

The details of laws regarding different groups can differ across jurisdictions, but the implications are generally similar. Most legislation differentiates between **direct discrimination** (treating people differently because of the group they belong to; this is almost universally outlawed and this is not something that any high-quality assessment is designed to do) and **indirect discrimination** (where a requirement or condition imposed has differential or adverse impact on different groups).

In the case of indirect discrimination against one or more group(s), the consequences can depend on whether any imposed condition that has been claimed to result in indirect discrimination can be justified. In terms of tests, requiring one group to complete a test and another not to without a clear justification (for example, having only

applicants who are not white take an English language test), or using different cut-off scores on a test for different groups, would not be appropriate.

The potential for indirect discrimination resulting from tests is something that test users need to consider carefully. If a test is used in selection and there are known and consistent average group performance differences, then the higher scoring group is more likely to pass a given cut score and will have a higher success rate. This could be indirect discrimination, unless the practice can be justified. Justification in this case could mean showing that the skill or characteristic measured was important for good performance on the job.

The general research literature shows that measures of cognitive ability typically reveal performance differences between certain groups of people. The use of job analysis and, where possible, local validation studies is particularly important for demonstrating the link between a test and the job it is being used to select for.

Research on the Saville aptitude tests shows differences which are in many cases less pronounced than has been found in other studies. Where differences have been found between different age, ethnicity and gender groups, these are generally small or moderate in size and broadly consistent with the established research literature.

For example, for Chinese candidates on a numerical reasoning test, in order to avoid the risk of indirect discrimination it is important to have evidence that the numerical reasoning skills measured by the test are an important skill for a substantial proportion of the job - and that people who do better at these skills are better at the job. It is also worth bearing in mind that it is often difficult to isolate single variables as the definite 'cause' of an apparent average group performance difference; for example, there is evidence that at least some ethnic group performance differences on certain cognitive tests may be due, at least in part, to differences in socioeconomic status and/or language proficiency.

It is thus especially important to bear in mind that any research carried out on group differences necessarily relies on a particular sample of individuals. Each sample of individuals is different and group differences should not be generalised beyond these specifically reported samples in an excessively broad manner.

In general, because of the predominantly small nature of group differences shown on the Saville aptitude tests, we do not advise that specific differences in profile interpretation should be warranted when considering test

results from different groups. Except where local legal frameworks permit or mandate such an approach, we do not recommend using separate norms for age, gender or ethnic groups.

A Note on Disability: Reasonable Adjustments and Special Accommodations

Many jurisdictions, including the UK, make legal provisions for individuals with disabilities and/or who require special accommodations. This can sometimes mean that reasonable adjustments are required during an assessment process to accommodate a person's specific requirements/disability in order to allow them as fair and comparable an assessment experience as possible. Accommodations may include allowing additional time, providing the assessment in another format (e.g. use of screen reading software or assistance by a sighted administrator) or providing a sign language interpreter for a hearing impaired person who communicates by signing.

Each organisation should take their own legal advice with regard to their human resource activities. Saville Assessment is not in a position to advise on legal matters.

Fair Assessment

Respect for the Individual and the Instrument

It can be beneficial to consider ethical considerations from the perspective of the candidate. As a trained test user, one can easily forget the concerns which may be felt by candidates. For the candidate, completing an assessment may be part of a life-changing experience (applying for a new job, developing competence, learning to work with a new team) whereas for the administrator it may not be a very interesting way to spend a day. Whatever an administrator feels about the process, an appreciation of its importance to the candidates and how this may affect them should drive the administrator's approach. The candidate may be nervous or frightened and this can affect their ability to understand instructions and comply with requests. Administrators need to deal with questions and problems in a patient and professional manner. Testing may be unfamiliar to candidates and they may be surprised by the formal nature of administration. Administrators should take this into account in explaining not just what will happen, but the purpose behind it. There is evidence that candidates are more likely to regard decisions as fair when they are aware of the processes used to reach the decision.

Choosing Appropriate Tests/Questionnaires

Tests and questionnaires should be chosen on the basis

of a thorough job analysis to ensure that decisions are being made on the basis of relevant information. To ensure assessment fairness, look for evidence of studies examining the appropriateness of the instrument with different groups. Ensure that administration is always conducted as specified to ensure that all candidates receive the same information and an equivalent assessment experience.

Preparing the Candidate

This is particularly important for aptitude tests. As discussed in the administration section, candidates should be briefed before completing psychometric instruments – on why the test or questionnaire is being used and what it involves, in addition to an explanation of what will happen with the data after collection. For example, it is usually recommended to advise candidates how their data will be used, how they will be stored and whether they may be used again in the future. Practice tests should always be provided wherever available.

Understanding what the test is about and how it will be used may help lessen a candidate's anxiety about the process. The unknown is often scarier. It will also allow a candidate to ask any questions or request special assistance for a disability, or any other reason. A proper briefing is also important so that a candidate understands the process. Only then can a candidate give their informed consent to participate.

Candidates are likely to be concerned about who will be able to see their results. This can be a particularly strong factor of concern for internal job applicants, or for those completing questionnaire about the culture of their organisation. Ensuring that candidates are fully aware of who may have access to their responses (and reassuring them about who will not) is a key element of ensuring that the candidate has provided their informed consent.

Professional Administration

The way in which the administration session is conducted can impact upon the performance and the views of the candidate regarding the questionnaire in use. Ensuring that candidates are given the administration instructions correctly and accurately is an important responsibility of the administrator, who should also take care to ensure that the candidates understand what they need to do and that they are completing the answer sheet correctly. An encouraging and professional attitude is required. The administrator should be approachable so that candidates are not frightened to ask questions, but should maintain a standardised and therefore reasonably formal atmosphere.

Dealing with Language

Language issues should also be taken into account. For

any psychometric measure you should consider what the impact of language needs are. Where English is not the primary language, consider whether it would be more appropriate to test in another language. How well does the candidate speak and read English? It is generally recommended to assess candidates in their language of greatest proficiency, wherever possible. If you are not sure of the implications for testing, take advice from someone with expertise in the area.

Dealing with Disability

Similarly, if a candidate has a disability, consider how this may affect the scores. A candidate with sight impairment may have difficulty reading a booklet or seeing a computer screen. Is there an alternative mode of administration that would help (e.g. using magnification equipment or a screen enlarger, or even having someone else read the content to them)? A candidate with dyslexia may also have difficulty reading some assessment content and may need more time than other candidates to complete the task. A candidate with a motor impairment may have difficulty using a mouse to fill in an answer sheet. As a general point, all of our standard Saville Assessment online aptitude tests use HTML5 technology, which is compatible with all modern computer and tablet browsers and permits a range of different adjustments.

It is important to discuss these issues with the candidate before the assessment to find out his or her needs. Often the candidates themselves have the greatest insight into how they could potentially be accommodated. In UK law, the concept of reasonable adjustment is important. Any adjustments to a process, including assessments, should be reasonable attempts to accommodate the specific individual's needs. The general principle is that any adjustment should attempt to provide the individual with a comparable assessment experience to other candidates, rather than attempting to deliberately make the process different. The assessor and assessee both have a responsibility for being as accommodating as is reasonable.

Where an accommodation requires changing the standardised administration procedure, you should always take advice from an expert regarding how this should be done and what issues will need to be taken into account. It may be that an accommodation may go beyond what you expect so it is important to seek expert advice rather than just relying on common sense views of what accommodations you personally consider reasonable.

Interpreting Scores

Care should always be taken to interpret an assessment

correctly. Be clear on what the scales measure – use the descriptions of the assessment provided in supporting materials to help you. Where you have adapted the test for a particular individual's needs, make sure you understand what impact this might have on the scores (if any) before interpreting them.

The choice of norm group will have an impact upon scores. Scores based on a very specific benchmark group, for example those based on people in a single profession, may need interpreting differently from those based on a broader norm. For example, an individual with prior sales experience and who is applying for a sales job may be seen to have a high score on a 'selling' scale when compared to a broad general population norm group which includes non-salespeople. The same person's score normed against a group of successful salespeople could be expected to result in a lower output score, because everyone in the group is higher on 'selling' compared to the majority of population. It is important to understand this difference and to ensure that scores are not misinterpreted as a result of the use of norm groups which are based on more specific samples of people, and hence where there is a narrower underlying spread of scores. Specific norm groups are also likely to have a smaller sample size than general population norm groups.

A similar approach can also be applied to mixed gender or ethnic group norms. As discussed above, worldwide research has shown certain differences between the average scores achieved by different groups on certain kinds of cognitive tests. A commonly reported difference between males and females is that women tend to score more highly on 'caring' scales. Thus, if a man and woman had an identical raw response on a caring scale, and were normed against single-sex groups only, the score of the man would appear much higher than that of the woman. If they were normed on a mixed gender group, the two would have an identical score. If one assumes that the 'caring' difference between men and women is a genuine one, using a mixed comparison group would seem more appropriate, as single-gender groups would artificially inflate the scores of males and deflate the scores of females.

What this does illustrate is the importance of using a norm group which is representative of the group being assessed. A norm based on predominantly young people, for example, would likely be less representative than one containing people of all ages. Choosing appropriate norm groups and understanding where there are patterns of scores associated with some demographic groups is an important part of interpreting scores. Users should be careful neither to stereotype individuals by their group membership.

Feedback

Candidates who have completed psychometric tests are likely to be very interested in their results. Giving candidates a 'right to reply', i.e. allowing them to comment on the results, not only provides a greater depth of information but also enables the candidate to provide explanations for responses and further information where appropriate. This is likely to make candidates feel more comfortable about the way in which their responses are being interpreted, particularly in high-stakes selection situations.

Test Use Policy

It is generally good practice for the use of tests to be guided by a test use policy. This will set out standards and local policies on a range of relevant issues. This helps ensure that minimum standards are maintained and that there is a consistency in practice across different assessment processes.

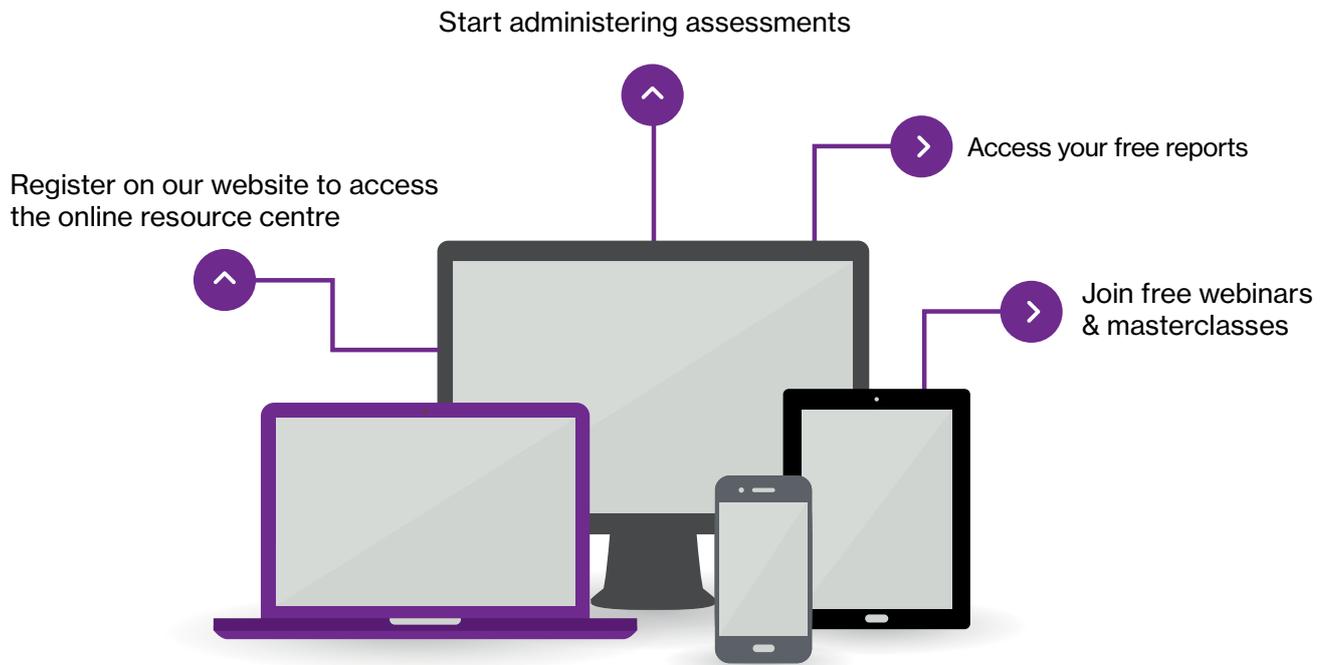
Training and Responsibility of Test Users

In order to be able to access certain kinds of assessments users must undergo training but, as with any set of skills or knowledge, over time issues may be forgotten and bad habits can develop. Equally, new developments may require updating of knowledge. It is important to engage in continuing professional development to maintain up-to-date knowledge and develop skills.

Finally, it is important for all test users to be aware of their responsibilities in following procedures and maintaining good practice. In the end, it is the responsibility of the test administrator to ensure proper practice. It is the responsibility of the test user to ensure that all interpretations from the test are valid and appropriate to the context and for the person who is using the information.

Notes:

Section 13: Next Steps



Contact our Duty Consultant for additional support.



+44 (0)208 619 9000



info@savilleassessment.com

Administering Tests: A Recap

- No annual subscription fees to use tests
- No separate administration, scoring or profiling fees
- **Saville Assessment Bureau – your managed service option:** Complete the form and our Bureau team will set up the questionnaire and send you reports
- **Saville Assessment Oasys – your self-service option:** One-off set up fee then a third off report costs. Your Oasys platform can be branded

Notes:

The Saville Assessment Community

What you can expect:

- Invitations to webinars and practitioner masterclasses
- Discussion on industry-specific news and topics
- Networking opportunities with other professionals
- Privileged offers



@savilleassess



www.linkedin.com/company/saville-assessment



www.facebook.com/SavilleAssess/



savilleassess

Notes

Resource Centre

As a trained user of Saville Assessment Products, you now have access to the Members Area of the Saville Assessment website. In this area you will find technical manuals and handbooks, product updates, research papers and survey results. To register, simply go to website www.savilleassessment.com and visit the Members Area. Please contact us if you require any assistance in registering.

Duty Consultant

The Duty Consultant is available Monday to Friday to share their specialist knowledge on Saville Assessment's product portfolio. The Duty Consultant provides advice on issues including product choice, appropriateness of comparison groups and interpretation of reports.

Administration

Our dedicated bureau service is open 8am to 6pm, Monday to Friday (except Bank Holidays) to set up your projects. In addition, the Bureau Service can offer support and advice for trained administrators managing their own Oasys platforms.

Straightforward Pricing

Online pricings include all aspects of the online Invited Access or Supervised Access aptitude and Work Strengths assessments, and receipt of the relevant report (excluding any consultancy support for face-to-face administration). There are no licence fees and, depending upon usage, a more cost-effective route is having a Saville Assessment Oasys platform, with a reduction in the cost per administration. The Oasys platform can be branded to reflect client corporate colours and logos. Please contact your course director for further information on Oasys.

Notes:

Section 14: Post-Course Work

Written Reports Overview

- BPS requirement:
 - Selection report for a hiring manager, summarising the test performance of two candidates
 - Feedback report for a candidate, summarising their performance across three tests
 - Produce a list of the things you would need to consider and do before, during and after a supervised test administration session
 - Write an email inviting a candidate to complete aptitude tests in an unsupervised setting
 - Set up a testing project using the Saville Assessment Bureau Request Form. Record a sample test session introduction and demonstrate your understanding of test invigilation. Hand mark and score an answer form.
- BPS fees to pay
- Complete within three months

Notes

Optional BPS Accreditation

If you would like to register with the British Psychological Society (BPS), you are required to complete some post-course work to ensure that you meet the requirements of the BPS. You can then be signed off by Saville Assessment as having met the required standards for the BPS Test User: Occupational Ability certificate.

Summary of Tasks

The requirements for accreditation with the BPS are outlined below.

- **Task 1: Selection Report** - Candidates A & B have completed Swift Analysis Aptitude as part of their recruitment process for the role of Business Analyst. Your task is to write a report to go to a hiring manager, summarising the candidates' test performance in relation to the requirements of the Business Analyst job description.
- **Task 2: Candidate Feedback Report** - Candidate C has completed Verbal Analysis Aptitude, Numerical Analysis Aptitude and Abstract Reasoning Aptitude as part of a recruitment process. Your task is to write a feedback report for Candidate C, summarising their performance on the tests.
- **Task 3: Test Administration** - Produce a list of the things you would need to consider and do before, during and after a supervised test administration session.
- **Task 4: Invitation Email** - Write an invitation email to a candidate, providing them with the relevant information they need before completing unsupervised tests.
- **Task 5: Test Administration and Scoring:**
 - 5a) Test Preparation: You have decided to progress Candidate A for the Business Analyst role based on their initial testing. As is best practice, you are now inviting this candidate to a supervised retest. Complete the Bureau Request form which will tell our support team how to set up a project to assess Candidate A. For the purpose of these tasks, Candidate A has now confirmed their attendance.
 - 5b) Test Introduction: You are administrating the session and need to audio record your introduction for the session.
 - 5c) Test Invigilation: Another candidate, Candidate D, indicated from the initial screening process that they had additional requirements and completed a hard-copy version of the assessment. Your task here is to invigilate this test session and record any deviant behaviours or issues that arise in the Test Session Record.
 - 5d) Test Marking: You need to mark Candidate D's hard-copy answer paper from their testing session. This kind of testing is now less common but it is illustrative of how raw scores are transformed into the standardised scores that help us to benchmark and compare candidates.

More information on the tasks can be found in this section.

Task 1: Selection Report on Two Candidates

Candidate A and Candidate B have applied for the role of Business Analyst (see job description on page 144) and have completed the Swift Analysis Aptitude test as part of their application. The candidates' results are on pages 145-159.

Your task is to complete one written report to go to the hiring manager, the IT Director, summarising the scores of both candidates in relation to the requirements of the Business Analyst job description.

To help you, please see the Example Written Selection Report on page 3. You can use this for guidance when writing your own report.

Guidance for writing your report:

- Highlight what each test is designed to measure
- Who will have access to these results and how long will the results typically be valid for?
- Explain the use of norm groups and relevance of the group
- Express each of the candidate's scores in terms which are meaningful, relative to the norm group, such as percentiles and performance bandings (e.g. low/average/high)
- Highlight where there are differences between the two candidates' performance on the total score
- It is important that you make clear that decisions should be made on the total score
- Include a summary and conclusion for the IT Director in your report, highlighting key points of consideration

Example Written Selection Report

Selection Report: Graduate Consultant

Overview

- This report gives a summary of two candidates' results on the online Verbal Analysis aptitude assessment.
- The test is used to decide which candidates should be progressed to the interview stage of the selection process for the role of Graduate Consultant.
- This report is for the attention of Graham Smith. It is confidential and should not be discussed with anyone other than those involved in the selection of candidates for the Graduate Consultant position.
- This report has been prepared specifically for the purpose stated and based on the information available.
- Psychometric reports are generally held to have a maximum period of validity of two years.
- On the basis of the test scores and the pre-established decision criteria, it is recommended that only Jane Moore is progressed to the interview stage.

Introduction

Previously, the candidates have passed the essential screening criteria for the Graduate Consultant role. At the current online aptitude testing stage, candidates need to perform better than 31% of the comparison group on the Verbal Analysis total score in order to progress to the interview stage.

The test used is designed for high level roles, and intended to assess those who have completed an undergraduate degree or are of graduate calibre. The test was used as research has shown that ability tests are powerful predictors of future workplace performance. In order to compare the candidates' performance to those of a similar educational background, their scores were compared to a group of 14,421 UK graduates. The nature of this comparative data should be considered when reading through the comments made in this report.

Verbal Analysis Assessment

The Verbal Analysis assessment covers the ability to understand written information and to draw appropriate conclusions from it. The test was selected following a thorough analysis of the job which revealed that a vital part of the role is to be able to understand and correctly interpret written information from clients or internal contacts. The test has been found to be reliable, consistently obtaining similar results from candidates, and valid, predicting workplace performance.

Candidate Scores

Jane Moore

Overall, Jane has performed better than 73% of the comparison group on the Verbal Analysis test, which is an above average score. Jane worked at a faster than average pace, responding to the questions more quickly than most people.

Stephanie McDonnell

Overall, Stephanie has performed better than 18% of the comparison group on the Verbal Analysis test, which is a below average score. Stephanie worked at a slower than average pace, responding to the questions more slowly than most people.

Summary

Jane had the highest performance overall on the test with an above average total score, compared to Stephanie's below average total score.

For further information on the assessment, please contact Mike Jones on extension 048.

Task 2: Candidate Feedback Report

Candidate C has completed the Saville Assessment Verbal Analysis, Numerical Analysis and Abstract Reasoning tests. Following a selection process for the role of Systems Administrator in which Candidate C has been informed that they are not to be progressed to a final interview stage, they were offered the opportunity for feedback. Candidate C has specifically requested full written feedback.

Your task is to write a feedback report for Candidate C. Candidate C's test scores are shown on pages 161-183.

To help you, please see the Example Written Candidate Feedback Report on page 4. You can use this for guidance when writing your own report.

Guidance for writing your report:

- Highlight what each test is designed to measure
- Who will have access to these results and how long the results will typically be valid for?
- Explain the use of norm groups and relevance of the group
- Express the individual's scores in terms which are meaningful, relative to the norm group, such as percentiles and performance bandings (e.g. low/average/high)
- We recommend using your own letterhead template for this report

Example Written Candidate Feedback Report

Feedback Report for Robin McLean

Overview

- This report gives a summary of your results on the Verbal Analysis and Numerical Analysis aptitude tests which you completed as part of the recent recruitment process for the role of Legal Advisor in CrossTech LLP.
- This report is your confidential copy. You are responsible for its safe-keeping and can decide who else will have access.
- This report has been prepared specifically for the purpose stated and based on the information available.
- Psychometric reports are generally held to have a maximum period of validity of two years.

Introduction

You completed the Verbal Analysis and Numerical Analysis aptitude assessments for selection for the role of Legal Advisor. This report summarises performance on the two aptitude assessments.

The tests used are designed for use with graduates, professionals, managers and directors. The tests were used as research has shown that ability tests are powerful predictors of future workplace performance. In order to compare your performance to those of a similar background, your scores were compared to a mixed group of 10,511 professionals and managers in the UK. The nature of this comparative data should be considered when reading through the comments made in this report.

Verbal Analysis Aptitude Assessment

The Professional Verbal Analysis assessment covers the ability to understand verbal information and to draw appropriate conclusions from it. The test was selected as a vital part of the Legal Advisor role is the ability to understand and interpret written information. The test has been found to be reliable, consistently obtaining similar results from candidates, and valid, predicting workplace performance.

Your Scores

Overall, you have performed better than 88% of the comparison group, which is an above average score. This indicates that you are likely to find working with verbal

information easier than many professionals and managers. When responding to the questions within the test, you worked at a faster pace than average.

Numerical Analysis Aptitude Assessment

The Professional Numerical Analysis assessment covers the ability to understand numerical information and to draw appropriate conclusions from it. The test was selected as a vital part of the Legal Advisor role is the ability to understand and interpret numerical information. The test has been found to be reliable, consistently obtaining similar results from candidates, and valid, predicting workplace performance.

Your Scores

Overall, you have performed better than 8% of the comparison group on the Numerical Analysis aptitude assessment, which is a below average score. This indicates that you are likely to find working with numerical information more difficult than many professionals and managers. When responding to the questions within the test, you worked at a much slower pace than average.

If you would like to develop your numerical analysis abilities, the following tips may be useful:

- When you read newspapers and reports pay attention to numerical information.
- Complete calculations both with and without a calculator.
- Look for differences such as percentage changes in numerical trends.
- Check calculations done by others.
- Take on responsibilities which involve working with numbers.

Summary

Your performance was above average on the Verbal Analysis Aptitude assessment and your performance on the Numerical Analysis Aptitude assessment was below average, indicating you may find working with numerical information more difficult than many people.

If you have any queries, please contact one of our team on extension 9035.

Task 3: Test Administration

Produce a list of the things you would need to consider and do before, during and after a supervised test administration session.

Task 4: Invitation Email

Write an invitation email to a candidate, providing them with the relevant information they need before completing a test in an unsupervised setting. An example email is provided in this section for guidance. Within your email, address the following points:

- a) How would you prepare candidates to do an online test?
- b) What would you do to ensure they have the necessary guidance and preparation?

Task 5: Test Administration and Scoring

5a – Test Preparation

You are inviting Candidate A to a supervised testing session following your decision to progress them ahead of Candidate B. Your task is to fill in the Bureau Request Form to set up the testing project for Candidate A's session. The Bureau Request Form lets our support team know which assessment you want to use, the norm group you want to benchmark against and any instructions for your delegates.

Please note, you are not actually submitting the request form, follow the instructions below to send us a preview link for your form.

Guidance for completing your request form:

- Use the Bureau Request Form link:
<https://insight2.towerswatson.com/wix/p30907064.aspx>
- Candidate A's details to include in this form are Candidate A;
sacandidatea@gmail.com
- In the first stage, Candidate A completed Swift Analysis Aptitude. You might now want to look in depth at their Verbal and Numerical ability, which are both important to the Business Analyst role (see job description on page 145 of your workbook)
- Select the same norm group that was used in the first stage of testing
- Include, "Level A post-course work" in your project name
- When you complete the Summary page of the form, please copy the link under Finish Later (highlighted in the image below). Send this link to us alongside your other post-course work when your tasks are ready to be marked:

- Please do not select, "Submit" on the Bureau Request Form, we only need to see the preview link you send to us.

5b – Test Introduction

Candidate A is attending your supervised training session. This session will use an online test so the test instructions and practice questions will be presented to the candidate on screen. Your task, as supervisor for this testing session, is to welcome Candidate A and introduce the session. Audio record yourself delivering this introduction.

Guidance for completing your audio test introduction:

- Use the template on page 48 of your workbook to support writing your introduction. Amend for the tests you have chosen to use for Candidate A.
- Consider the tone you use to put your candidate at ease for the testing session.
- Save your audio file as a WAV, MP3, WMA file or similar, and attach this to your email alongside the other coursework tasks you submit.

Saville Assessment
Willis Towers Watson

Summary

Project Details	
REQUESTED BY Example Example@example.com	COMPANY Example
BILLING CONTACT (NAME) Example	PO NUMBER N/A
BILLING CONTACT (TEL NO.) 0777777777	COST CENTRE N/A
BILLING CONTACT (EMAIL) Example@example.com	VAT NUMBER N/A
BILLING CONTACT (ADDRESS) Example	DELIVERY ADDRESS Same as billing
PROJECT NAME Example	PROJECT TYPE NO RESPONSE
FEEDBACK PROVIDER Example Example@example.com	OTHER REPORT RECIPIENTS N/A

Note
By completing and submitting this form, you confirm that you accept the standard terms and conditions which can be viewed at:
<https://www.savilleassessment.com/terms-conditions>. You also confirm that you have the appropriate qualifications to administer and interpret the assessments you have requested.

Finish Later
<https://insight2.towerswatson.com/wix/p30907064.aspx?r=25423&e=1TXWOCMNI9>

Need Help?
If you have any questions or require assistance completing this form, please send an email to Saville Assessment Bureau.

5c - Test Invigilation

Historically, testing was only conducted with pencil and paper assessments. To prepare for Candidate A's testing session, watch this video that outlines a traditional test administration session. In a real testing session you would need to complete a Test Session Record to track the session and have a written account of anything you want to flag to the hiring manager.

1. Practice filling in the Test Session Record (page 9 in this document) as you watch this video. You may not be able to fill all of the boxes but should complete as far as you can.

Example of pencil and paper testing

<https://www.youtube.com/watch?v=IH1-PJnnpPQ&feature=youtu.be>

2. Explain how you would manage each, "deviant behaviour" listed below. These are actual deviant behaviours candidates might exhibit in a real testing session. Please also include reasoning for your actions. For example, a candidate asks, "When do we get feedback?" after the test. How you manage this will be dependent on your hiring process and your company's testing and feedback policy but you are able to give the candidate this information.

Deviant Behaviours

Before the test...

- Starting before the test administrator has said it is okay to do so
- The candidate asks, "Who will see the results?"

During the test...

- The candidates asks, "How long do we have left?"
- The candidate asks, "Can I leave a question out?"

After the test...

- The candidate says, "I've done this test before..."
- The candidate asks, "Can you explain question five to me?"

Test Session Record

Venue	
Test Administrator Name	
Test Session Time	

Test					
Version					
Test Time (Length)					
Test Session Began (Clock Time)					
Test Session Finished (Clock Time)					
No. of Candidates					

Any interruptions, event, disturbances	
Any deviant behaviours	
Any questions or issues raised by candidates	

5d – Test Marking

Candidate D indicated that they had additional requirements which necessitated their use of a hard-copy assessment. Your task is to mark Candidate D's completed Verbal Reasoning answer form on page 12.

The instructions and examples below will help explain how to calculate their total score, speed, accuracy and caution and then report the Percentile, Sten and T-score range for Candidate D's total score.

- Follow each answer across to the relevant shaded box. Indicate each correct answer with a ✓ and each incorrect answer with a x in that box. If an answer is blank, leave the relevant box blank (see example below; in Figure 1, question 4, the correct answer is C but no response is given; the corresponding box is therefore left blank).
- Total the correct (✓), incorrect (x) and blank answers for each column and check that the sum of these matches the number of items (see Figure 2).
- Now total each of the columns to get the number of correct (✓), incorrect (x) and blank answers.
- Use the Blank (BLA), Attempted (ATT), and Total (TOT) figures to calculate the Accuracy (ACC), Speed (SPE) and Caution (CAU) values (see Figure 2).
- Transfer all figures to the Score Summary Table, (see Figure 3).
- Use the Profile Chart to work out and report the percentile, Sten and T-score range. You need to identify the total score on the Profile Chart and follow this down to the %ile, Stens, T-scores box at the bottom of the page. Identifying the corresponding %ile, Sten and T-score is a way of transforming your raw score into a normed score. This means you can benchmark Candidate D's responses against an external group. For the example in Figure 4, the candidate's total score is 17 which is at the 46th percentile, Sten 5 and their T-Score range is 45 – 49.
- You need to send through the total score, accuracy, speed, caution and the percentile, Sten and T-score range for Candidate D's total score. You do not need to report percentile, Sten and T-score range for any of Candidate D's other scores (i.e their Accuracy, Speed, Caution or the types of questions completed).

Figure 1.

Figure 2.

	✓	X	Blanks	No. of Items
UWM	2	2	1	5
CT	3	1	2	6
MVI	4	1	0	5
EWM	6	1	0	7
CA	2	1	2	5
TOTAL	T17	6	B5A	28

28	-	BLA 5	=	ATT 23
TOT 17	/	ATT 23	=	ACC .74
ATT 23	/	SPE 28	=	SPE .82
ACC .74	-	SPE .82	=	CAU -.08

Figure 3.

Score Summary Table:		Score	%ile	Sten	T
TOTAL SCORE (TOT)	17	46	5	45-49	
Accuracy (ACC)	.74				
Speed (SPE)	.82	You do not need			
Caution (CAU)	-.08	to report the			
Understanding Word Meaning (UWM)	2	Percentile, Sten			
Comprehending Text (CT)	3	or T for the			
Making Verbal Inferences (MVI)	4	remaining scores			
Evaluating Written Materials (EWM)	6	for this task.			
Comparing Arguments (CA)	2				

Figure 4.

Example Profile Chart

Scores		Low			Below Average			Average			Above Average			High										
Total	TOTALSCORE(TOT)	<=5	7	9	11	13	15	17	19	21	23	25	27											
		6	8	10	12	14	16	18	20	22	24	26	28											
Test Taking Style	Accuracy(ACC)	<=.32	.36	.39	.43	.46	.49	.53	.56	.60	.63	.66	.70	.73	.77	.80	.83	.87	.90	.94	.97	1		
		.34	.37	.41	.44	.48	.51	.54	.58	.61	.65	.68	.71	.75	.78	.82	.85	.88	.92	.95	.99			
	Speed (SPE)	<=.61		.68		.75		.82		.89		.96												
		.64		.71		.79		.86		.93		1												
	Caution(CAU)	<= -.61	-.58	-.54	-.50	-.46	-.43	-.39	-.35	-.31	-.28	-.24	-.20	-.17	-.13	-.09	-.05	-.02	.02	.06	.09	.13	.17	>=.21
		-.59	-.56	-.52	-.48	-.45	-.41	-.37	-.33	-.30	-.26	-.22	-.18	-.15	-.11	-.07	-.04	-.00	.04	.08	.11	.15	.19	
Item Type	Understanding Word Meaning (UWM)	<=1					3					5												
		2							4															
	Comprehending Text (CT)	0				2				4											6			
		1							3															
	Making Verbal Inferences (MVI)	0				2					4													
		1							3													5		
Evaluating Written Materials (EWM)	0					2				4										6				
	1				1				3															7
Comparing Arguments (CA)	0					2					4													
	1								3															5
%ile		1	2	4	5	8	12	16	21	27	34	42	50	58	66	73	79	84	88	92	95	96	98	99
		2	3	4	7	10	14	18	24	31	38	46	54	62	69	76	82	86	90	93	96	97	98	
Stens		1	2	3	4	5	6	7	8	9	10													
T - scores		25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64	65 - 69	70 - 74													

Candidate D Answer Form

	A	B	C	D
1	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
3	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	A	B	C	D
5	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
6	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
7	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	A	B	C	D
9	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
11	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
12	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

	A	B	C	D
13	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
14	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
15	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
16	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	A	B	C	D
17	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
18	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
19	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
20	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

	A	B	C	D
21	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
23	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

	A	B	C	D
25	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
27	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

	UWM	CT	MVI	EWM	CA
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					

Correct answer indicator
 Answered
 Both answers must be filled in for item to be correct

Score Summary

	✓	X	Blanks	No. of Items
UWM				5
CT				6
MVI				5
EWM				7
CA				5
TOTAL	TOT		BLA	28

$28 - \text{BLA} = \text{ATT}$
 $\text{TOT} / \text{ATT} = \text{ACC}^*$
 $\text{ATT} / 28 = \text{SPE}^*$
 $\text{ACC} - \text{SPE} = \text{CAU}$

* Round to 2 decimal places

Profile Chart

Scores		Low			Below Average			Average			Above Average			High													
Total	TOTAL SCORE(TOT)	<=5	7	9	11	13	15	17	19	21	23	25	27														
		6	8	10	12	14	16	18	20	22	24	26	28														
Test Taking Style	Accuracy(ACC)	<=.32	.36	.39	.43	.46	.49	.53	.56	.60	.63	.66	.70	.73	.77	.80	.83	.87	.90	.94	.97	1					
		.34	.37	.41	.44	.48	.51	.54	.58	.61	.65	.68	.71	.75	.78	.82	.85	.88	.92	.95	.99						
	Speed(SPE)	<=.61		.68		.75		.82		.89		.96															
			.64		.71		.79		.86		.93		1														
	Caution(CAU)	<= -.61	-.58	-.54	-.50	-.46	-.43	-.39	-.35	-.31	-.28	-.24	-.20	-.17	-.13	-.09	-.05	-.02	.02	.06	.09	.13	.17	>=.21			
			-.59	-.56	-.52	-.48	-.45	-.41	-.37	-.33	-.30	-.26	-.22	-.18	-.15	-.11	-.07	-.04	-.00	.04	.08	.11	.15	.19			
Item Type	Understanding Word Meaning(UWM)	<=1					3										5										
			2								4																
	Comprehending Text(CT)	0					2					4										6					
				1						3							5										
	Making Verbal Inferences(MVI)	0					2						4														
		1								3											5						
Evaluating Written Materials(EWM)		0					2					4										6					
						1				3							5								7		
Comparing Arguments(CA)	0						2						4														
											3																
	%ile	1	2	4	5	8	12	16	21	27	34	42	50	58	66	73	79	84	88	92	95	96	98	99			
		2	3	4	7	10	14	18	24	31	38	46	54	62	69	76	82	86	90	93	96	97	98				
	Stens	1	2	3	4	5	6	7	8	9	10																
	T - scores	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64	65 - 69	70 - 74																

Next Steps:

When you have completed the reports, lists and invite email, email/post them to your Course Director:

Saville Assessment
CI Tower
1st Floor St George's Square
New Malden
KT3 4HG

Tel: 020 8619 9000

Email: training.administrator@savilleassessment.com

Job Description

Role: Business Analyst

Report to: IT Director

Tradigital are looking for a self-motivated team player to join our fast growing business and to support our Account Management Team in providing leading e-learning and application technology solutions to clients.

We have an exciting opportunity for a recent graduate with strong analytical and IT skills to join our solutions-focused business, providing excellent service for our global clients. The Business Analyst will act as a link between our information technology team and client facing team, providing robust analysis of business needs and client requirements. The ideal candidate will be someone who can provide quality solutions and analysis to complex IT problems within short timescales.

Key Responsibilities:

- Prepare and deliver weekly update presentations to the business
- Work closely and communicate regularly with the Account Management Teams to understand client requirements
- Analysing written and numerical data to inform business and client solutions
- Produce written reports on analysis results and support the Account Management Team in writing technical client proposals
- Providing quality solutions for client IT problems
- Assist in writing and preparing business plans
- Improving and testing current services and new IT solutions
- Identify system faults and the causes
- Implementing new system processes

Required Skills and Experience:

- Educated to degree level, or two years' relevant experience
- Excellent written and verbal communication skills
- Excellent numerical and IT skills (including Excel)
- Strong problem solving skills
- Ability to multi-task and plan effectively
- Strong attention to detail

Desirable Skills:

- Knowledge of e-learning authoring tools and learning management systems
- Experience of working with and designing IT systems and processes



Assessment Report Candidate A



Swift Analysis
Aptitude-Rx

Contents

Introduction to Assessment Report.....	3
Total Score.....	4
Aptitude & Pace Information.....	5
Improving Abilities.....	6
Online Test Access Summary (For Assessor Use).....	7

About this Report

This report is based upon Swift Analysis Aptitude, an online test of the ability to reason with information presented in verbal, numerical and diagrammatic formats.

The results are compared against a group of 10,511 professionals and managers in the UK. The results in this report are presented on a 1 to 10 Sten scale, where 1 indicates low performance and 10 indicates high performance on the test. The margin of error that should be allowed before concluding that there is a difference between scores is indicated by the diamond shape.

When reading this report, please remember that it is based on the information gained from the test completion only. It describes performance on this particular test, rather than performance at work or study. Research suggests that ability tests can be powerful predictors of successful performance in study and work activities.

The information contained in this report is confidential and every effort should be made to ensure that it is stored in a secure place.

The information contained within this report is likely to provide a valid measure of aptitude for 12 to 24 months.

The report is based on the results of the online test that the respondent was invited to complete under unsupervised conditions. The identity of the actual respondent has not been verified by a test administrator. Further testing under supervised conditions is recommended for high-stake decision making.

This report was produced using Saville Assessment software systems and has been generated electronically. Saville Assessment do not guarantee that it has not been changed or edited. We can accept no liability for the consequences of the use of this report.

The application of this test is limited to Saville Assessment employees, agents of Saville Assessment and clients authorised by Saville Assessment.

Introduction to Assessment Report

This report provides feedback on the responses of Candidate A to the Swift Analysis Aptitude test.

Total Score

This test measures verbal, numerical and diagrammatic analysis, which are important in the world of work for a variety of roles. This section of the report provides a total test score relative to the comparison group: Professionals & Managers (UK; IA; 2015)

The Total Score indicates how well Candidate A has performed overall on the test.

Aptitude Area Sub-Scores

The sub-scores provide information on how Candidate A performed on each of the aptitude sub-tests. The pattern of results indicates relative strengths and weaknesses across the following areas of aptitude:

Verbal - assesses the ability to understand, interpret and evaluate written information.

Numerical - assesses the ability to understand, interpret and evaluate numerical data.

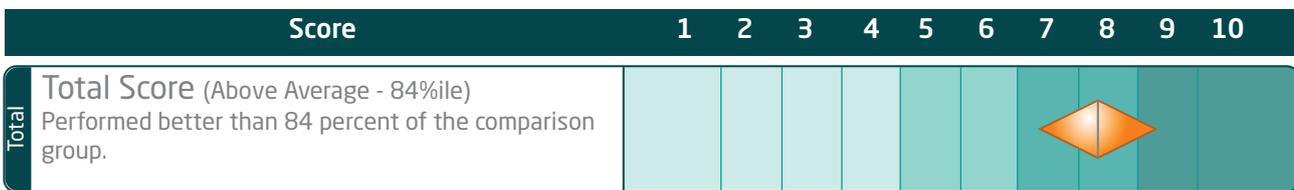
Diagrammatic - assesses the ability to analyse diagrams, sequences and transformations.

Aptitude & Pace Comparison

Aptitude and pace scores are shown for each of the areas in the test. These scores are compared in a graph using a 1 to 10 sten scale, with the sten values given in brackets. The pace score is based on the candidate's response time for the questions they completed compared to the average response time for the same questions. Pace is shown from slow at the bottom of the graph to fast at the top. Aptitude runs from low on the left of the graph to high on the right.

Total Score

This page shows the Total Score relative to the Professionals & Managers (UK; IA; 2015) comparison group on a 1 to 10 sten scale.



Interpretation Guidelines

Comparison Group: Professionals & Managers (UK; IA; 2015)

Sten 1: higher than about 1% of the comparison group

Sten 2: higher than about 5% of the comparison group

Sten 3: higher than about 10% of the comparison group

Sten 4: higher than about 25% of the comparison group

Sten 5: higher than about 40% of the comparison group

Sten 6: higher than about 60% of the comparison group

Sten 7: higher than about 75% of the comparison group

Sten 8: higher than about 90% of the comparison group

Sten 9: higher than about 95% of the comparison group

Sten 10: higher than about 99% of the comparison group

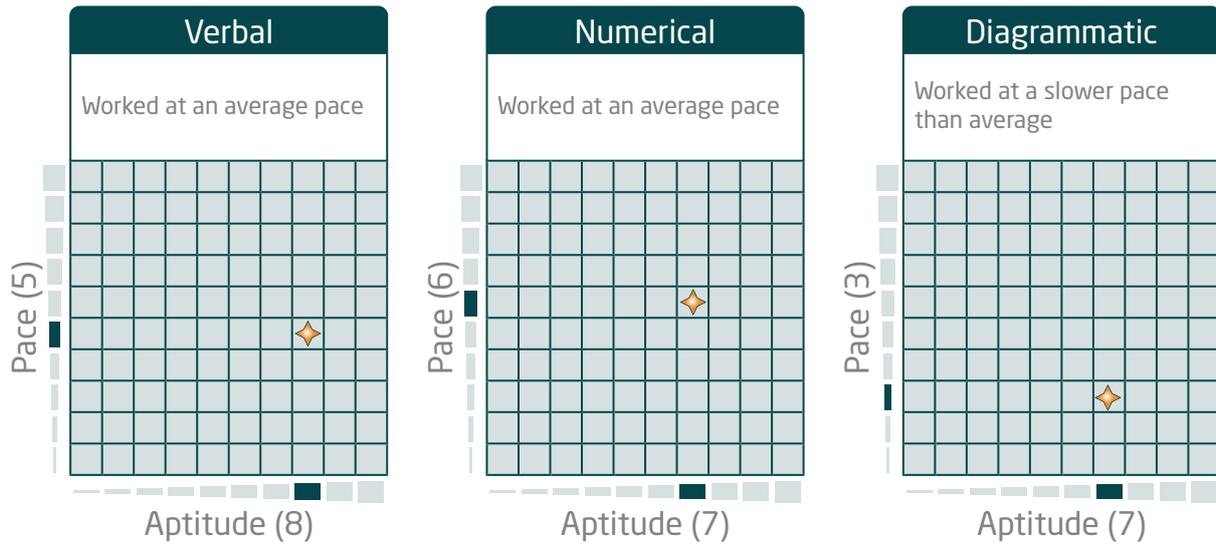
Aptitude & Pace Information

This page displays aptitude and pace information for each of the areas in the test relative to the Professionals & Managers (UK; IA; 2015) comparison group.

Aptitude Area Sub-Scores

Scores		1	2	3	4	5	6	7	8	9	10
Aptitude Areas	Verbal (Above Average - 86%ile) Likely to find working with verbal information easier than other people.										
	Numerical (Above Average - 69%ile) Likely to find working with numerical information easier than other people.										
	Diagrammatic (Above Average - 73%ile) Likely to find working with diagrammatic information easier than other people.										

Aptitude & Pace Comparison



Improving Abilities

Some tips for improving abilities are provided below:

Verbal

- When you read newspapers and articles, try to establish the main points.
- Look up the meaning of unfamiliar words.
- Read passages of text and pick out the key details.
- Compare written arguments, looking for similarities and differences between them.
- Look at something you have written and rewrite it more concisely.

Numerical

- When you read newspapers and reports, pay attention to numerical information.
- Complete calculations both with and without a calculator.
- Look for differences such as percentage changes in numerical trends.
- Check calculations done by others.
- Take on responsibilities which involve working with numbers.

Diagrammatic

- Examine diagrams in books and newspapers.
- Study flowcharts of processes and procedures.
- Improve your logic by solving puzzles.
- Try to clarify different types of relationships within diagrams.
- Create diagrams in order to illustrate sequences of events.

Online Test Access Summary (For Assessor Use)

This section of the report provides additional information about the test completion.

Initial Access: 01/06/2018 (17:49 GMT)
Responses Saved: 01/06/2018 (18:28 GMT)
Administrator Resets: 0
Candidate Aborts: 1
Time Adjustment: None

Assessment Report Candidate B



Swift Analysis
Aptitude-Rx

Contents

Introduction to Assessment Report.....	3
Total Score.....	4
Aptitude & Pace Information.....	5
Improving Abilities.....	6
Online Test Access Summary (For Assessor Use).....	7

About this Report

This report is based upon Swift Analysis Aptitude, an online test of the ability to reason with information presented in verbal, numerical and diagrammatic formats.

The results are compared against a group of 10,511 professionals and managers in the UK. The results in this report are presented on a 1 to 10 Sten scale, where 1 indicates low performance and 10 indicates high performance on the test. The margin of error that should be allowed before concluding that there is a difference between scores is indicated by the diamond shape.

When reading this report, please remember that it is based on the information gained from the test completion only. It describes performance on this particular test, rather than performance at work or study. Research suggests that ability tests can be powerful predictors of successful performance in study and work activities.

The information contained in this report is confidential and every effort should be made to ensure that it is stored in a secure place.

The information contained within this report is likely to provide a valid measure of aptitude for 12 to 24 months.

The report is based on the results of the online test that the respondent was invited to complete under unsupervised conditions. The identity of the actual respondent has not been verified by a test administrator. Further testing under supervised conditions is recommended for high-stake decision making.

This report was produced using Saville Assessment software systems and has been generated electronically. Saville Assessment do not guarantee that it has not been changed or edited. We can accept no liability for the consequences of the use of this report.

The application of this test is limited to Saville Assessment employees, agents of Saville Assessment and clients authorised by Saville Assessment.

Introduction to Assessment Report

This report provides feedback on the responses of Candidate B to the Swift Analysis Aptitude test.

Total Score

This test measures verbal, numerical and diagrammatic analysis, which are important in the world of work for a variety of roles. This section of the report provides a total test score relative to the comparison group: Professionals & Managers (UK; IA; 2015)

The Total Score indicates how well Candidate B has performed overall on the test.

Aptitude Area Sub-Scores

The sub-scores provide information on how Candidate B performed on each of the aptitude sub-tests. The pattern of results indicates relative strengths and weaknesses across the following areas of aptitude:

Verbal - assesses the ability to understand, interpret and evaluate written information.

Numerical - assesses the ability to understand, interpret and evaluate numerical data.

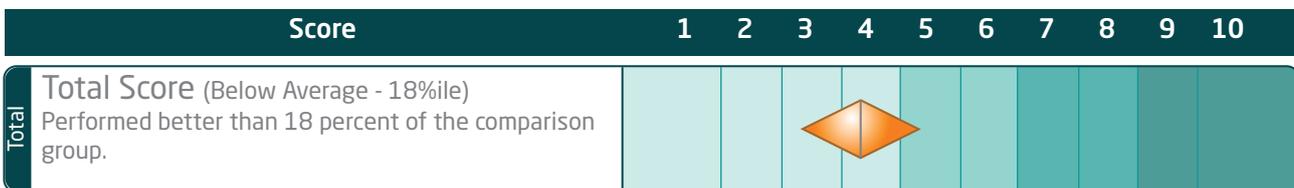
Diagrammatic - assesses the ability to analyse diagrams, sequences and transformations.

Aptitude & Pace Comparison

Aptitude and pace scores are shown for each of the areas in the test. These scores are compared in a graph using a 1 to 10 sten scale, with the sten values given in brackets. The pace score is based on the candidate's response time for the questions they completed compared to the average response time for the same questions. Pace is shown from slow at the bottom of the graph to fast at the top. Aptitude runs from low on the left of the graph to high on the right.

Total Score

This page shows the Total Score relative to the Professionals & Managers (UK; IA; 2015) group on a 1 to 10 sten scale.



Interpretation Guidelines

Comparison Group: Professionals & Managers (UK; IA; 2015)

Sten 1: higher than about 1% of the comparison group

Sten 2: higher than about 5% of the comparison group

Sten 3: higher than about 10% of the comparison group

Sten 4: higher than about 25% of the comparison group

Sten 5: higher than about 40% of the comparison group

Sten 6: higher than about 60% of the comparison group

Sten 7: higher than about 75% of the comparison group

Sten 8: higher than about 90% of the comparison group

Sten 9: higher than about 95% of the comparison group

Sten 10: higher than about 99% of the comparison group

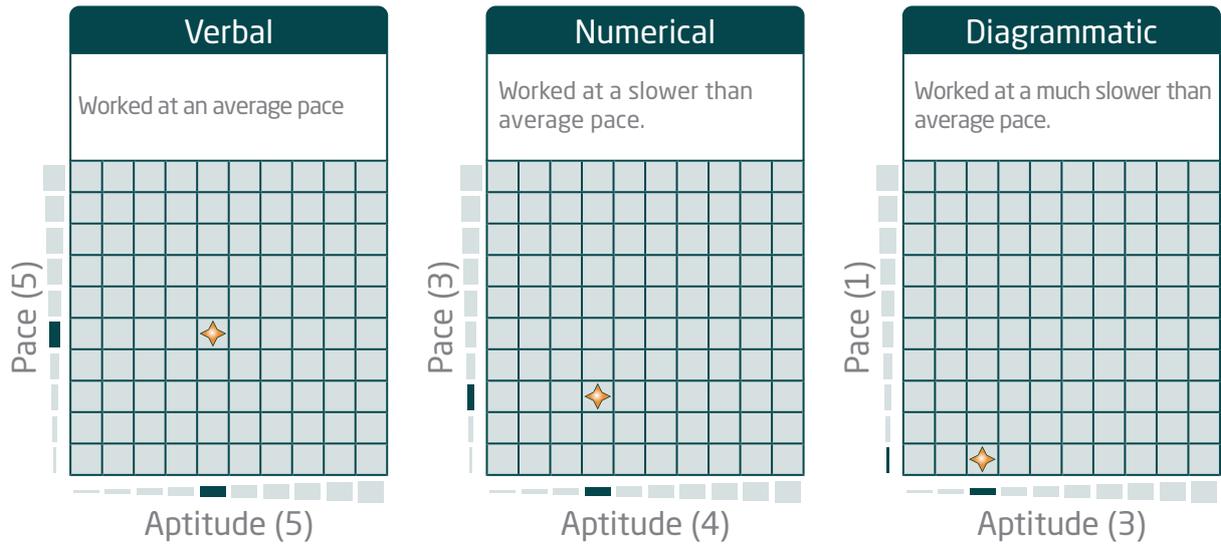
Aptitude & Pace Information

This page displays aptitude and pace information for each of the areas in the test relative to the Professionals & Managers (UK; IA; 2015) comparison group.

Aptitude Area Sub-Scores

Scores		1	2	3	4	5	6	7	8	9	10
Aptitude Areas	Verbal (Average - 34%ile) Likely to find working with verbal information as easy as other people.					◆					
	Numerical (Below Average - 27%ile) Likely to find working with numerical information more difficult than other people.				◆						
	Diagrammatic (Below Average - 14%ile) Likely to find working with diagrammatic information more difficult than other people.			◆							

Aptitude & Pace Comparison



Improving Abilities

Some tips for improving abilities are provided below:

Verbal

- When you read newspapers and articles, try to establish the main points.
- Look up the meaning of unfamiliar words.
- Read passages of text and pick out the key details.
- Compare written arguments, looking for similarities and differences between them.
- Look at something you have written and rewrite it more concisely.

Numerical

- When you read newspapers and reports, pay attention to numerical information.
- Complete calculations both with and without a calculator.
- Look for differences such as percentage changes in numerical trends.
- Check calculations done by others.
- Take on responsibilities which involve working with numbers.

Diagrammatic

- Examine diagrams in books and newspapers.
- Study flowcharts of processes and procedures.
- Improve your logic by solving puzzles.
- Try to clarify different types of relationships within diagrams.
- Create diagrams in order to illustrate sequences of events.

Online Test Access Summary (For Assessor Use)

This section of the report provides additional information about the test completion.

Initial Access: 01/06/2018 (20:57 GMT)
Responses Saved: 01/06/2018 (23:21 GMT)
Administrator Resets: 0
Candidate Aborts: 0
Time Adjustment: None

Assessment Report
Candidate C

Abstract Reasoning
Aptitude-Rx

Contents

Introduction to Assessment Report.....	3
Total Score.....	4
Aptitude & Pace Information.....	5
Improving Abilities.....	6
Online Test Access Summary (For Assessor Use).....	7

About this Report

This report is based upon Abstract Reasoning Aptitude, an online test of the ability to reason with information presented in an abstract format.

The results are compared against a group of 2,510 professionals and managers in the UK. The results in this report are presented on a 1 to 10 Sten scale, where 1 indicates low performance and 10 indicates high performance on the test. The margin of error that should be allowed before concluding that there is a difference between scores is indicated by the diamond shape.

When reading this report, please remember that it is based on the information gained from the test completion only. It describes performance on this particular test, rather than performance at work or study. Research suggests that ability tests can be powerful predictors of successful performance in study and work activities.

The information contained in this report is confidential and every effort should be made to ensure that it is stored in a secure place.

The information contained within this report is likely to provide a valid measure of aptitude for 12 to 24 months.

The report is based on the results of the online test that the respondent was invited to complete under unsupervised conditions. The identity of the actual respondent has not been verified by a test administrator. Further testing under supervised conditions is recommended for high-stake decision making.

This report was produced using Saville Assessment software systems and has been generated electronically. Saville Assessment do not guarantee that it has not been changed or edited. We can accept no liability for the consequences of the use of this report.

The application of this test is limited to Saville Assessment employees, agents of Saville Assessment and clients authorised by Saville Assessment.

Introduction to Assessment Report

This report provides feedback on the responses of Candidate C to the Abstract Reasoning Aptitude test.

Total Score

This test measures abstract reasoning, which is important in the world of work for a variety of roles. This section of the report provides a total test score relative to the comparison group: Professionals & Managers (UK; IA; 2015)

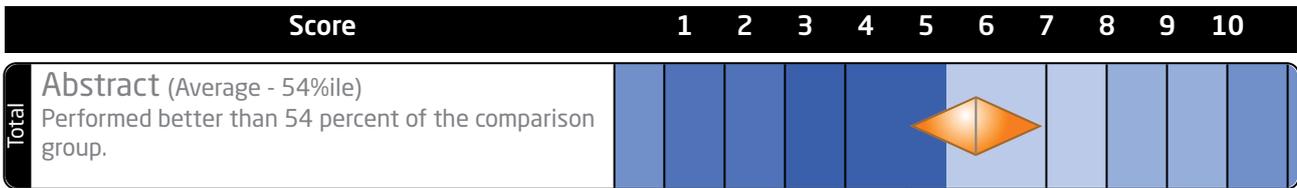
The Total Score indicates how well Candidate C has performed overall on the test.

Aptitude & Pace Comparison

Aptitude and pace scores are shown for each of the areas in the test. These scores are compared in a graph using a 1 to 10 sten scale, with the sten values given in brackets. The pace score is based on the candidate's response time for the questions they completed compared to the average response time for the same questions. Pace is shown from slow at the bottom of the graph to fast at the top. Aptitude runs from low on the left of the graph to high on the right.

Total Score

This page shows the Total Score relative to the Professionals & Managers (UK; IA; 2015) comparison group on a 1 to 10 sten scale.



Interpretation Guidelines

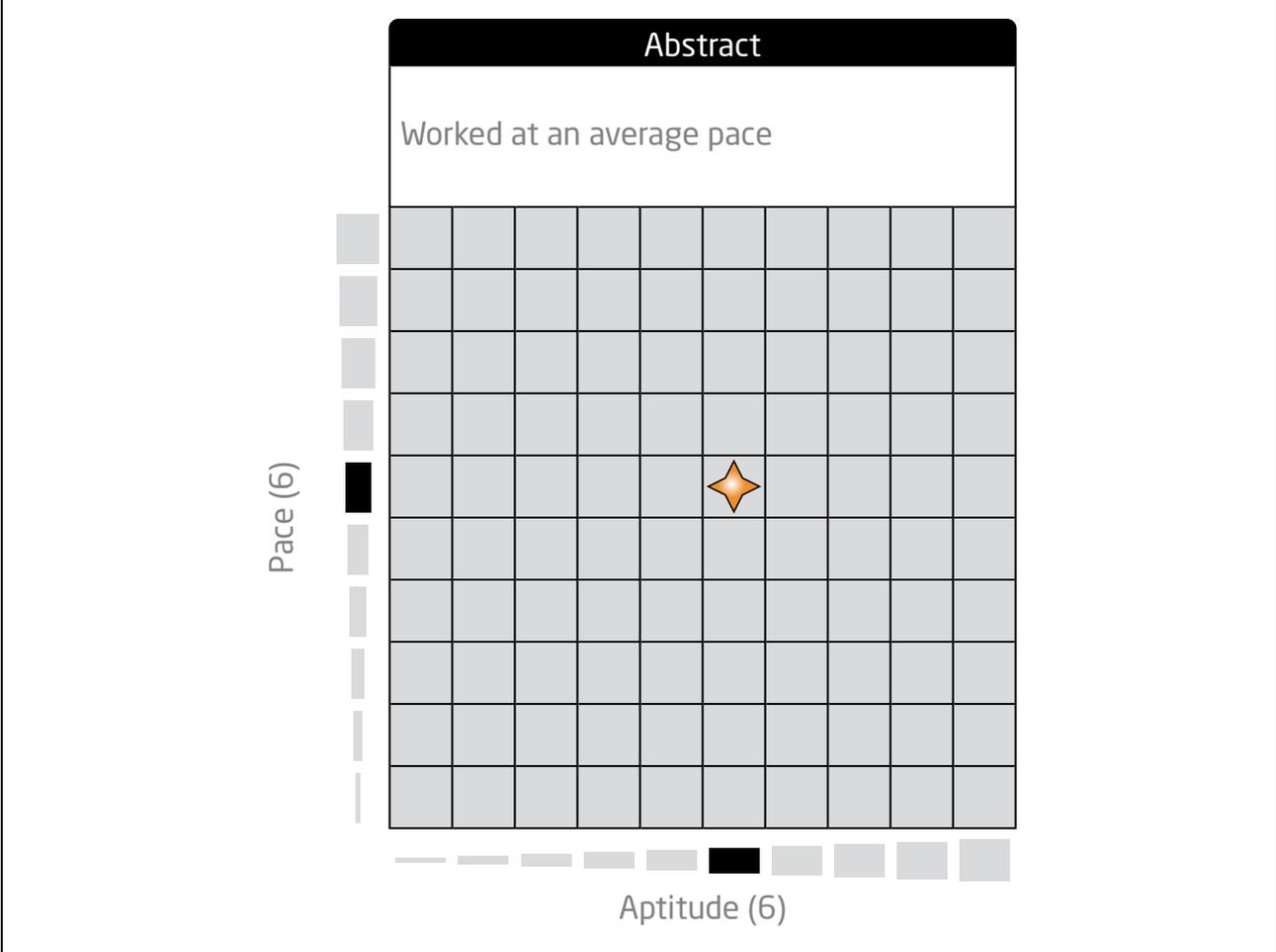
Comparison Group: Professionals & Managers (UK; IA; 2015)

- Sten 1: higher than about 1% of the comparison group
- Sten 2: higher than about 5% of the comparison group
- Sten 3: higher than about 10% of the comparison group
- Sten 4: higher than about 25% of the comparison group
- Sten 5: higher than about 40% of the comparison group
- Sten 6: higher than about 60% of the comparison group
- Sten 7: higher than about 75% of the comparison group
- Sten 8: higher than about 90% of the comparison group
- Sten 9: higher than about 95% of the comparison group
- Sten 10: higher than about 99% of the comparison group

Aptitude & Pace Information

This page displays aptitude and pace information for the test relative to the Professionals & Managers (UK; IA; 2015) comparison group.

Aptitude & Pace Comparison



Improving Abilities

Some tips for improving abilities are provided below:

Abstract

- Work with abstract materials.
- Examine information presented in abstract forms in books and newspapers.
- Complete logic puzzles and games.
- Look for patterns and relationships in information.
- Practise creating diagrams which represent relationships, connections and sequences.

Online Test Access Summary (For Assessor Use)

This section of the report provides additional information about the test completion.

Initial Access: 01/06/2018 (11:12 GMT)
Responses Saved: 01/06/2018 (11:31 GMT)
Administrator Resets: 0
Candidate Aborts: 0
Time Adjustment: None

Assessment Report Candidate C



Numerical Analysis

Aptitude-Rx

Contents

Introduction to Assessment Report.....	3
Total Score.....	4
Aptitude & Pace Information.....	5
Improving Abilities.....	6
Online Test Access Summary (For Assessor Use).....	7

About this Report

This report is based upon Numerical Analysis Aptitude, an online test of the ability to reason with information presented in a numerical format.

The results are compared against a group of 10,511 professionals and managers in the UK. The results in this report are presented on a 1 to 10 Sten scale, where 1 indicates low performance and 10 indicates high performance on the test. The margin of error that should be allowed before concluding that there is a difference between scores is indicated by the diamond shape.

When reading this report, please remember that it is based on the information gained from the test completion only. It describes performance on this particular test, rather than performance at work or study. Research suggests that ability tests can be powerful predictors of successful performance in study and work activities.

The information contained in this report is confidential and every effort should be made to ensure that it is stored in a secure place.

The information contained within this report is likely to provide a valid measure of aptitude for 12 to 24 months.

The report is based on the results of the online test that the respondent was invited to complete under unsupervised conditions. The identity of the actual respondent has not been verified by a test administrator. Further testing under supervised conditions is recommended for high-stake decision making.

This report was produced using Saville Assessment software systems and has been generated electronically. Saville Assessment do not guarantee that it has not been changed or edited. We can accept no liability for the consequences of the use of this report.

The application of this test is limited to Saville Assessment employees, agents of Saville Assessment and clients authorised by Saville Assessment.

Introduction to Assessment Report

This report provides feedback on the responses of Candidate C to the Numerical Analysis Aptitude test.

Total Score

This test measures numerical analysis, which is important in the world of work for a variety of roles. This section of the report provides a total test score relative to the comparison group: Professionals & Managers (UK; IA; 2015)

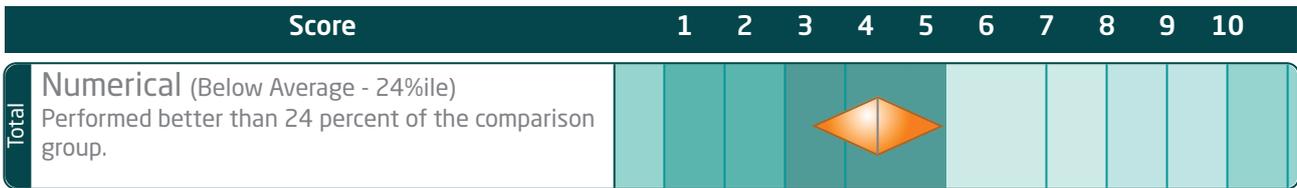
The Total Score indicates how well Candidate C has performed overall on the test.

Aptitude & Pace Comparison

Aptitude and pace scores are shown for each of the areas in the test. These scores are compared in a graph using a 1 to 10 sten scale, with the sten values given in brackets. The pace score is based on the candidate's response time for the questions they completed compared to the average response time for the same questions. Pace is shown from slow at the bottom of the graph to fast at the top. Aptitude runs from low on the left of the graph to high on the right.

Total Score

This page shows the Total Score relative to the Professionals & Managers (UK; IA; 2015) comparison group on a 1 to 10 sten scale.



Interpretation Guidelines

Comparison Group: Professionals & Managers (UK; IA; 2015)

Sten 1: higher than about 1% of the comparison group

Sten 2: higher than about 5% of the comparison group

Sten 3: higher than about 10% of the comparison group

Sten 4: higher than about 25% of the comparison group

Sten 5: higher than about 40% of the comparison group

Sten 6: higher than about 60% of the comparison group

Sten 7: higher than about 75% of the comparison group

Sten 8: higher than about 90% of the comparison group

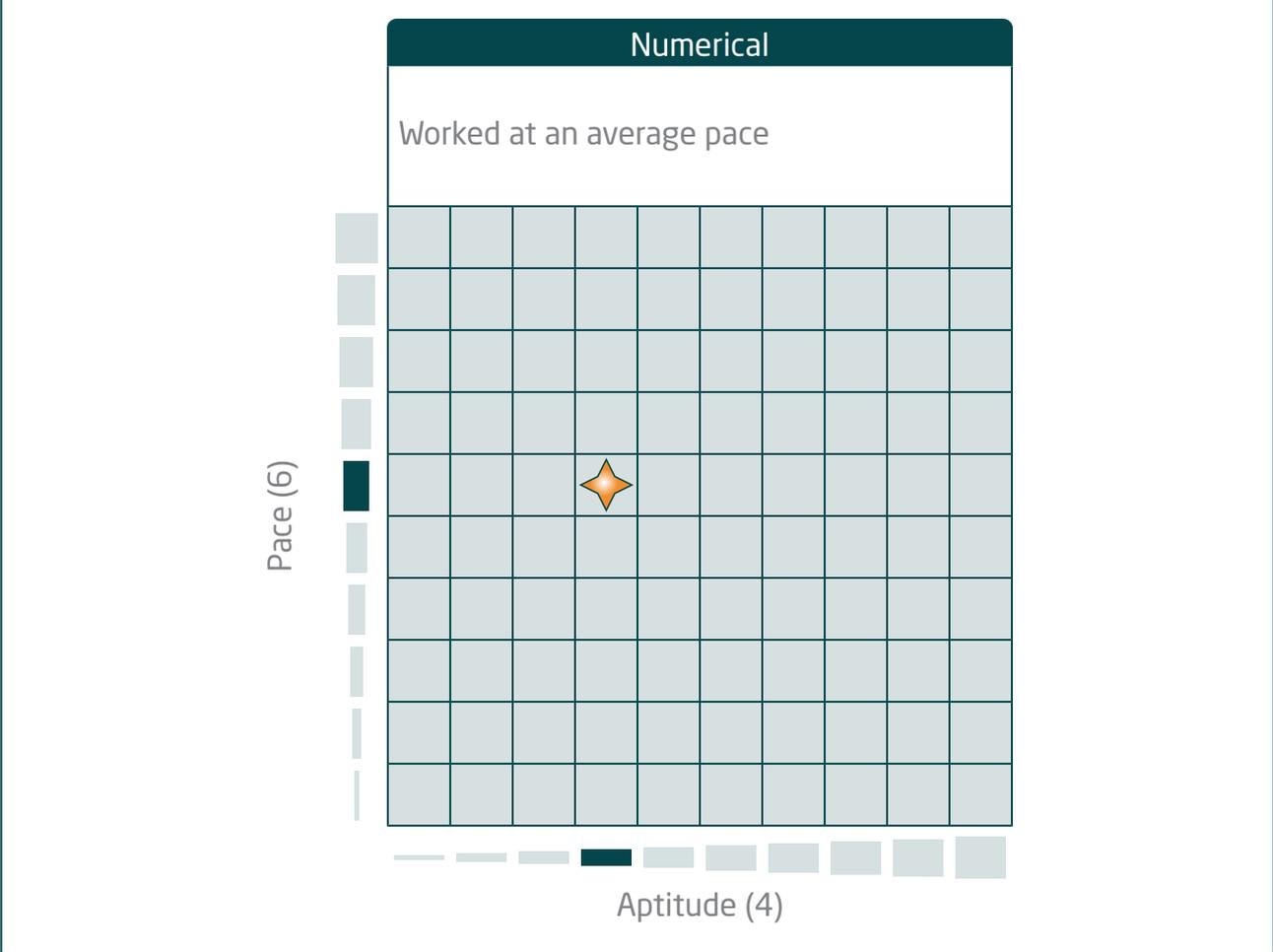
Sten 9: higher than about 95% of the comparison group

Sten 10: higher than about 99% of the comparison group

Aptitude & Pace Information

This page displays aptitude and pace information for the test relative to the Professionals & Managers (UK; IA; 2015) comparison group.

Aptitude & Pace Comparison



Improving Abilities

Some tips for improving abilities are provided below:

Numerical

- When you read newspapers and reports, pay attention to numerical information.
- Complete calculations both with and without a calculator.
- Look for differences such as percentage changes in numerical trends.
- Check calculations done by others.
- Take on responsibilities which involve working with numbers.

Online Test Access Summary (For Assessor Use)

This section of the report provides additional information about the test completion.

Initial Access: 01/06/2018 (09:10 GMT)
Responses Saved: 01/06/2018 (09:35 GMT)
Administrator Resets: 0
Candidate Aborts: 0
Time Adjustment: None

Assessment Report Candidate C



Verbal Analysis

Aptitude-Rx

Contents

Introduction to Assessment Report.....	3
Total Score.....	4
Aptitude & Pace Information.....	5
Improving Abilities.....	6
Online Test Access Summary (For Assessor Use).....	7

About this Report

This report is based upon Verbal Analysis Aptitude, an online test of the ability to reason with information presented in a verbal format.

The results are compared against a group of 10,511 professionals and managers in the UK. The results in this report are presented on a 1 to 10 Sten scale, where 1 indicates low performance and 10 indicates high performance on the test. The margin of error that should be allowed before concluding that there is a difference between scores is indicated by the diamond shape.

When reading this report, please remember that it is based on the information gained from the test completion only. It describes performance on this particular test, rather than performance at work or study. Research suggests that ability tests can be powerful predictors of successful performance in study and work activities.

The information contained in this report is confidential and every effort should be made to ensure that it is stored in a secure place.

The information contained within this report is likely to provide a valid measure of aptitude for 12 to 24 months.

The report is based on the results of the online test that the respondent was invited to complete under unsupervised conditions. The identity of the actual respondent has not been verified by a test administrator. Further testing under supervised conditions is recommended for high-stake decision making.

This report was produced using Saville Assessment software systems and has been generated electronically. Saville Assessment do not guarantee that it has not been changed or edited. We can accept no liability for the consequences of the use of this report.

The application of this test is limited to Saville Assessment employees, agents of Saville Assessment and clients authorised by Saville Assessment.

Introduction to Assessment Report

This report provides feedback on the responses of Candidate C to the Verbal Analysis Aptitude test.

Total Score

This test measures verbal analysis, which is important in the world of work for a variety of roles. This section of the report provides a total test score relative to the comparison group: Professionals & Managers (UK; IA; 2015)

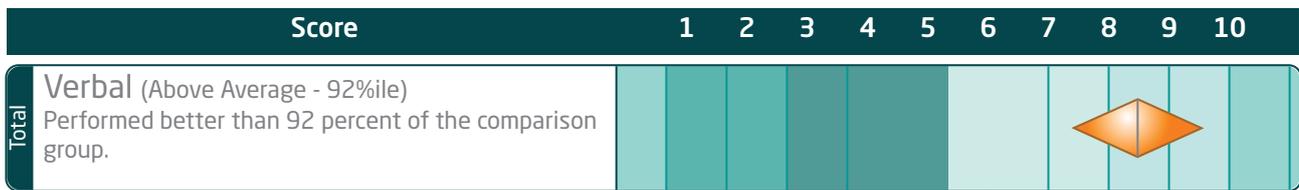
The Total Score indicates how well Candidate C has performed overall on the test.

Aptitude & Pace Comparison

Aptitude and pace scores are shown for each of the areas in the test. These scores are compared in a graph using a 1 to 10 sten scale, with the sten values given in brackets. The pace score is based on the candidate's response time for the questions they completed compared to the average response time for the same questions. Pace is shown from slow at the bottom of the graph to fast at the top. Aptitude runs from low on the left of the graph to high on the right.

Total Score

This page shows the Total Score relative to the Professionals & Managers (UK; IA; 2015) comparison group on a 1 to 10 sten scale.



Interpretation Guidelines

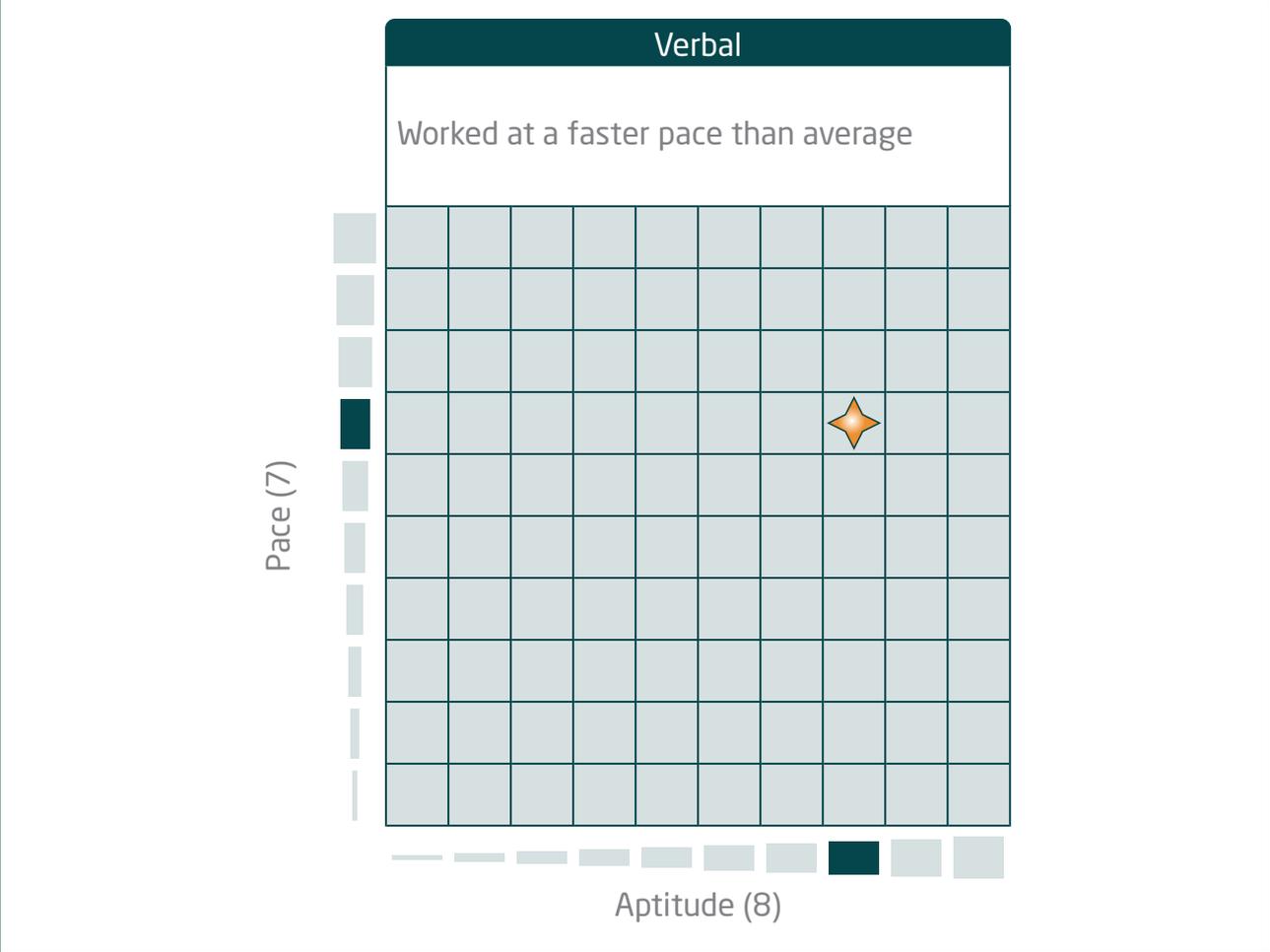
Comparison Group: Professionals & Managers (UK; IA; 2015)

- Sten 1: higher than about 1% of the comparison group
- Sten 2: higher than about 5% of the comparison group
- Sten 3: higher than about 10% of the comparison group
- Sten 4: higher than about 25% of the comparison group
- Sten 5: higher than about 40% of the comparison group
- Sten 6: higher than about 60% of the comparison group
- Sten 7: higher than about 75% of the comparison group
- Sten 8: higher than about 90% of the comparison group
- Sten 9: higher than about 95% of the comparison group
- Sten 10: higher than about 99% of the comparison group

Aptitude & Pace Information

This page displays aptitude and pace information for the test relative to the Professionals & Managers (UK; IA; 2015) comparison group.

Aptitude & Pace Comparison



Improving Abilities

Some tips for improving abilities are provided below:

Verbal

- When you read newspapers and articles, try to establish the main points.
- Look up the meaning of unfamiliar words.
- Read passages of text and pick out the key details.
- Compare written arguments, looking for similarities and differences between them.
- Look at something you have written and rewrite it more concisely.

Online Test Access Summary (For Assessor Use)

This section of the report provides additional information about the test completion.

Initial Access: 01/06/2018 (09:41 GMT)
Responses Saved: 01/06/2018 (10:02 GMT)
Administrator Resets: 0
Candidate Aborts: 0
Time Adjustment: None

Example Written Selection Report

Selection Report: Graduate Consultant

Overview

- This report gives a summary of two candidates' results on the online Verbal Analysis aptitude assessment.
- The test is used to decide which candidates should be progressed to the interview stage of the selection process for the role of Graduate Consultant.
- This report is for the attention of Graham Smith. It is confidential and should not be discussed with anyone other than those involved in the selection of candidates for the Graduate Consultant position.
- This report has been prepared specifically for the purpose stated and based on the information available.
- Psychometric reports are generally held to have a maximum period of validity of two years.
- On the basis of the test scores and the pre-established decision criteria, it is recommended that only Jane Moore is progressed to the interview stage.

Introduction

Previously, the candidates have passed the essential screening criteria for the Graduate Consultant role. At the current online aptitude testing stage, candidates need to perform better than 31% of the comparison group on the Verbal Analysis total score in order to progress to the interview stage.

The test used is designed for high level roles, and intended to assess those who have completed an undergraduate degree or are of graduate calibre. The test was used as research has shown that ability tests are powerful predictors of future workplace performance. In order to compare the candidates' performance to those of a similar educational background, their scores were compared to a group of 14,421 UK graduates. The nature of this comparative data should be considered when reading through the comments made in this report.

Verbal Analysis Assessment

The Verbal Analysis assessment covers the ability to understand written information and to draw appropriate conclusions from it. The test was selected following a thorough analysis of the job which revealed that a vital part of the role is to be able to understand and correctly interpret written information from clients or internal contacts. The test has been found to be reliable, consistently obtaining similar results from candidates, and valid, predicting workplace performance.

Candidate Scores

Jane Moore

Overall, Jane has performed better than 73% of the comparison group on the Verbal Analysis test, which is an above average score. Jane worked at a faster than average pace, responding to the questions more quickly than most people.

Stephanie McDonnell

Overall, Stephanie has performed better than 18% of the comparison group on the Verbal Analysis test, which is a below average score. Stephanie worked at a slower than average pace, responding to the questions more slowly than most people.

Summary

Jane had the highest performance overall on the test with an above average total score, compared to Stephanie's below average total score.

For further information on the assessment, please contact Mike Jones on extension 048.

Example Written Candidate Feedback Report

Feedback Report for Robin McLean

Overview

- This report gives a summary of your results on the Verbal Analysis and Numerical Analysis aptitude tests which you completed as part of the recent recruitment process for the role of Legal Advisor in CrossTech LLP.
- This report is your confidential copy. You are responsible for its safe-keeping and can decide who else will have access.
- This report has been prepared specifically for the purpose stated and based on the information available.
- Psychometric reports are generally held to have a maximum period of validity of two years.

Introduction

You completed the Verbal Analysis and Numerical Analysis aptitude assessments for selection for the role of Legal Advisor. This report summarises performance on the two aptitude assessments.

The tests used are designed for use with graduates, professionals, managers and directors. The tests were used as research has shown that ability tests are powerful predictors of future workplace performance. In order to compare your performance to those of a similar background, your scores were compared to a mixed group of 10,511 professionals and managers in the UK. The nature of this comparative data should be considered when reading through the comments made in this report.

Verbal Analysis Aptitude Assessment

The Professional Verbal Analysis assessment covers the ability to understand verbal information and to draw appropriate conclusions from it. The test was selected as a vital part of the Legal Advisor role is the ability to understand and interpret written information. The test has been found to be reliable, consistently obtaining similar results from candidates, and valid, predicting workplace performance.

Your Scores

Overall, you have performed better than 88% of the comparison group, which is an above average score. This indicates that you are likely to find working with verbal

information easier than many professionals and managers. When responding to the questions within the test, you worked at a faster pace than average.

Numerical Analysis Aptitude Assessment

The Professional Numerical Analysis assessment covers the ability to understand numerical information and to draw appropriate conclusions from it. The test was selected as a vital part of the Legal Advisor role is the ability to understand and interpret numerical information. The test has been found to be reliable, consistently obtaining similar results from candidates, and valid, predicting workplace performance.

Your Scores

Overall, you have performed better than 8% of the comparison group on the Numerical Analysis aptitude assessment, which is a below average score. This indicates that you are likely to find working with numerical information more difficult than many professionals and managers. When responding to the questions within the test, you worked at a much slower pace than average.

If you would like to develop your numerical analysis abilities, the following tips may be useful:

- When you read newspapers and reports pay attention to numerical information.
- Complete calculations both with and without a calculator.
- Look for differences such as percentage changes in numerical trends.
- Check calculations done by others.
- Take on responsibilities which involve working with numbers.

Summary

Your performance was above average on the Verbal Analysis Aptitude assessment and your performance on the Numerical Analysis Aptitude assessment was below average, indicating you may find working with numerical information more difficult than many people.

If you have any queries, please contact one of our team on extension 9035.

Example Aptitude Assessment Invitation Email

To:	Timothy.Webster@gmail.com
CC:	recruitment@GreenworthLeisure.com
Subject:	Greenworth Leisure Branch Manager Role - Manchester
Attachments:	Preparation guides.doc

Dear Timothy,

I am pleased to inform you that following your application for the above position, we would like to invite you to complete two aptitude tests – one on Verbal Analysis and one on Numerical Analysis. These assessments are being used as such skills are important in the job you have applied for. The tests are timed, each lasting exactly 20 minutes. You should allow an additional 15 minutes (approx.) per test for instructions and example questions to be completed.

You are required to complete the assessments by Friday 18th November. You will need internet access to complete the assessments. You will be sent the link to the assessments and some unique login details to this email address: **Timothy.Webster@gmail.com**. Please let me know as soon as possible if you foresee any difficulties accessing the assessments.

Please note that you may be required to complete further verbal and numerical aptitude assessments under supervised conditions later in the process.

I have included Preparation Guides for each of these assessments for you to work through. They explain the types of questions asked and give you the opportunity to attempt some practice questions. The real test questions start at a similar difficulty level but get progressively more difficult.

Please contact me if you have any special conditions that may impact on the assessments, including dyslexia.

The information you provide will be stored confidentially and will only be available to you and to those involved in the recruitment process, and in accordance with any applicable privacy policy notified to you. The data will not be used for any other purpose and will be stored for 12 months. All data will be securely stored in line with applicable legislation. You will be able to contact me for feedback on your performance on the aptitude assessments.

Please contact me if you have any specific queries or would like to know more about the ability assessments before completing them.

Yours sincerely,

Tina Smart

Recruitment Manager

Greenworth Leisure

Section 15: Review Worksheets

Section - Introduction to Testing

Q1. Distinguish between tests assessing 'maximum' performance and those assessing 'typical' performance.

Q2. List one key feature of each of the following types of workplace assessment:

Work Sample Tests

Aptitude Tests

Interest Inventories

Personality Assessments

Q3. What is the difference between each of the following methods of assessing someone's 'Can Do' ability? Give an example of each.

IQ tests

Aptitude Tests

Attainment tests

Section 2 - Job Analysis

Q1. Explain the purpose of job analysis and competency profiling.

Section 4 - Test Administration

Q1. Distinguish between Open, Controlled, Managed & Supervised modes of administration.

Q2. Describe the main issues associated with administering different types of tests in each mode.

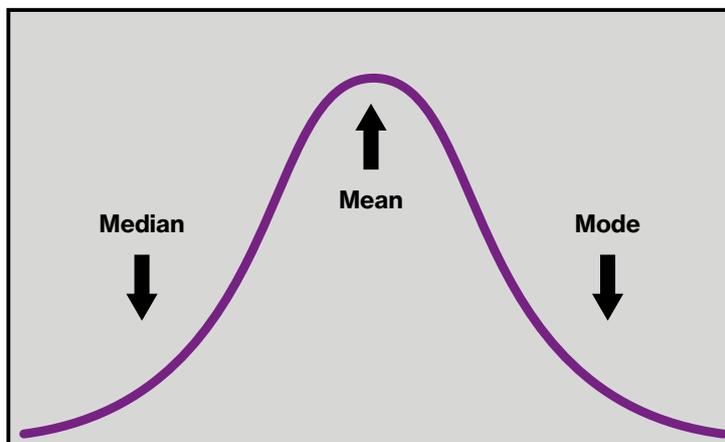
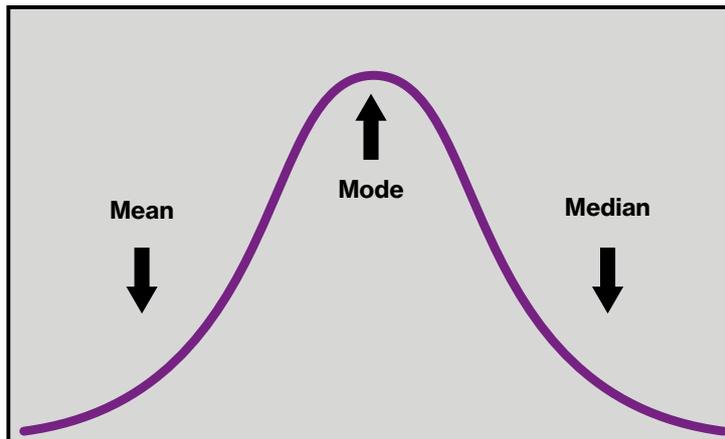
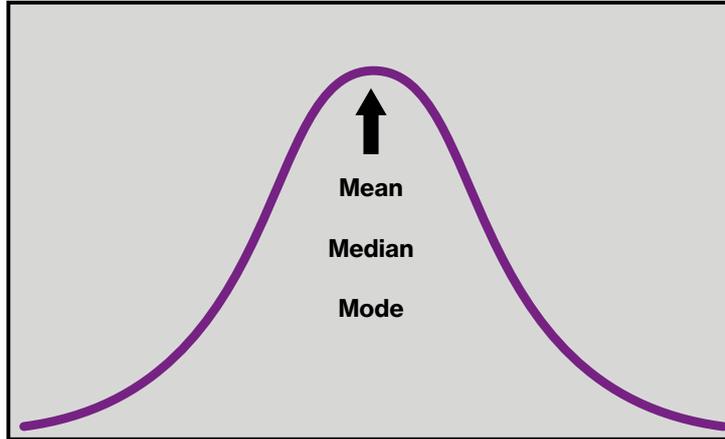
Q3. How can you overcome the potential problems of cheating with Invited Access (Controlled mode) administration?

Q4. Where would you find information on when to reset candidates who are 'locked out' of their test?

Q5. If a candidate asks to have access to the actual test prior to completion, what should you do?

Section 7 - Test Norms

Q1. Select which graph indicates the correct position of the mean, median and mode.



Q2. What is the difference between a raw score and a standardised score? Give an example of each.

Q3. What are the differences between a raw score, banding and ranking (e.g. percentiles) of candidates? What are the advantages and disadvantages of each?

Q4. What factors do you need to consider when getting norm samples?

Q5. How does the standard error of the mean change as the sample size increases?

Section 8 - Correlation

Q1. Describe the conditions under which a correlation is maximised and minimised.

Section 9 - Reliability

Q1. What can be sources of error in tests?

Q2. What is the basic principle of 'Classical Test Theory'?

Q3. How does the length of a test affect its reliability?

Q4. Describe what is meant by reliability and why it is important for measurement.

Q5. How are the main types of reliability assessed?

Q6. What are the advantages and drawbacks of the main types of test reliability?

Q7. What factors will influence the reliability estimates of a sample?

Q8. A candidate's raw score on a verbal reasoning test is 20 out of 28. In order to understand the possible error around this score, you have been asked to manually calculate the standard error of measurement.

$$SEm = SD \sqrt{1 - r_{tt}}$$

SEm = Standard error of measurement

SD = Standard deviation of scores for the reliability sample

r_{tt} = the reliability coefficient (test-retest, alternate form or internal consistency)

(a) Calculate the standard error of measurement. The reliability of the test is 0.84 and the standard deviation is 5.51 raw scores.

(b) What would be the upper and lower raw score values for the 68% confidence interval?

(c) What does this range represent?

Section 10 - Validity

Q1. What is the validity of a test essentially concerned with and how is it measured?

Q2. What is the difference between test reliability and test validity? How does one impact on the other?

Q3. What type of validity is concerned with whether a test looks reasonable to a candidate or line manager?

Q4. What form of validity evidence is often described as an overarching or universal concept of validity?

Q5. How is concurrent validity assessed?

Q6. How is predictive validity assessed?

Q7. How are test validity coefficients affected by job performance measures with low reliability?

Q8. What is meta-analysis? How does this technique overcome the sampling error problems of small studies?

Q9. What is the validity for ability tests from meta-analytic research?

Q10. What factors impact on test utility?

Section 11 - Approaches to Online Testing

Q1. Describe the basic concept of item response theory.

Q2. What does item response theory allow to happen when different candidates see different content?

Q3. Describe what is meant by a test information function.

Q4. Describe the principle of adaptive testing.

Section 12 - Best Practice & Ethics

Q1. You are recruiting for a line operative in a manufacturing plant. On reviewing your test data for the last two years, you notice a small difference in the performance of older women on a test of spatial reasoning; they are performing less well as a group on average than others.

a) What possible reasons can you think of for this difference in performance on the test?

b) What actions could you consider to ensure that you are recruiting a similar proportion of older women as other groups?

Q2. Give examples of environmental factors that could impact on an individual's test scores over the longer-term.

Q3. In high stakes recruitment what are the implications of using:

a) Mixed gender or ethnic group vs. single gender or ethnic group?

b) Occupationally related vs. general population norms?

Q4. What is a 'reasonable adjustment'?

Q5. Give examples of accommodations to the testing process which could be made for disabled candidates.

Q6. What issues need to be considered when interpreting adjusted scores?

Q7. Name two key legal acts which should be considered in assessment.

Q8. Referring to the funnel you created in the Assessment Choice section, if 90% of candidates were screened out at the invited access test stage, what would be the likely impact on:

a) The utility of the test?

b) The percentage of female applicants in the later stages of the assessment process?

c) The percentage of black applicants in the later stages of the assessment process?

Section 16: Resources

Test Session Record

Organisation: _____

Date: _____

Venue:	
Test Administrator Names:	
Test Session Time:	

Test:					
Version:					
Test Time (Length)					
Test Session Began (Clock Time)					
Test Session Finished (Clock Time)					
No of Candidates					

Any interruptions, event, disturbances	
Any deviant behaviours	
Any questions or issues raised by candidates	

Sample Psychometric Test Use Policy

CrossTech LLP

Test Use Policy

This document is for the attention of Authorised Test Users

Produced by: Mike Jones, Test Use Policy Co-ordinator

This document is designed to set out CrossTech LLP's psychometric test use policy. At CrossTech LLP we use psychometric tests to provide a robust, fair and objective form of information for selection and development decisions. We are committed to the highest standards of practice in test use, to ensure both our organisations and test-takers benefit from the use of psychometric tests. CrossTech LLP prides itself on being an equal opportunities employer and we are committed to policies and processes which promote diversity and inclusion. Throughout this policy, the guidance is designed to select and develop individuals on merit, fairly, equitably and objectively.

This document sets out how this is achieved by the appropriate use of tests. Please contact Mike Jones, the Test Use Policy Co-ordinator with any questions relating to this document.

1.0 Test Use

The use, administration, interpretation and feedback of psychometric tests is restricted to authorised test users: those with the relevant Test User: Occupational Ability qualification. Psychometric tests may also be administered by those with the relevant Assistant Test User qualification.

All use of psychometric assessments should comply with the policies outlined in this document and any local statutory and regulatory requirements. Tests should only ever be used for occupational assessment purposes (e.g. selection and development).

1.1 Test Choice

- Only authorised, qualified test users should make choices on which tests and norms to use, and this should be based on thorough job analysis and understanding of the person specification.
- It is the test user's responsibility to select an appropriate test which is reliable, valid and fair.
- Authorised test users are responsible for how the tests are used to support any decision making, e.g. establishing a cut-off score, weighting information relative to other assessment data collected, etc.

1.2 Preparation

- Preparation materials should be provided, wherever possible, to test-takers prior to the completion of an assessment.
- Live psychometric tests which are used for decision making should not be used to prepare or coach candidates for the purpose of being more effective on live psychometric tests. Practice tests are designed for this purpose.

1.3 Administration

- Assessments described as "Invited Access" (IA) are provided for completion in an online, unsupervised environment, e.g. on tablet, PC, laptop etc.
- Authorised test users are recommended to use invited access tests where the content is drawn from large banks of randomised question content.
- Where invited access tests are used to screen candidates, given the identity of the test-taker is not certain, a second test should be administered to candidates under supervised conditions (where identity can be verified) later in the selection process but before a job offer is made. The authorised test user is responsible for advising on any decision rules based on the scores in the second test which would disqualify the candidate from appointment to the given role.
- Where a supervised assessment is used, candidate identity should be checked in line with our general policy on recruitment with a valid form of photographic identification and proof of address.
- Tests which are for supervised use should be administered and interpreted by trained personnel according to the instructions provided.

1.4 Resets - Invited Access Assessments

- As a security precaution in some online, unsupervised tests (e.g. Saville Swift assessments), candidates will be 'locked out' of the assessment under certain circumstances, e.g. navigating away from the test browser.
- In such circumstances, a reset may be required to allow the candidate to complete the test.
- Our reset policy is to allow one reset per candidate without questioning the reasons for the 'lock-out'. If a second reset is requested, the authorised user should contact the candidate to discuss the situation and make an informed decision about whether another reset is permitted or not.
- Multiple candidate resets may start to provide a candidate with an advantage. Where a candidate is requiring multiple resets the recommendation is to require the candidate to complete the assessment under supervised conditions.

1.5 Languages

- Tests should be selected on the basis of their availability in the expected first language(s) of the candidate group.
- Where possible, test-takers should be tested in their first language to provide them with the best opportunity to demonstrate their full aptitude.
- Where language proficiency is a requirement for the

role, this should be tested with a language proficiency assessment, not a psychometric aptitude test.

1.6 Feedback: Right of request for information in a meaningful form

- Following completion of a test, it is good practice to offer feedback to test-takers. Feedback should be given by an authorised test user and should be given in a format and style which is appropriate and meaningful for the recipient.
- When test-takers are given feedback, this should be provided in compliance with any applicable statutory or regulatory requirements and the individual should be informed when and how the feedback will be delivered.
- The individual should be given the right to reply, including an opportunity to seek clarification and/or further explanation with respect to the feedback provided.

1.7 Validation

- Where group sizes allow, i.e. 100+ job incumbents have completed a test for a particular role, and there is good quality performance data, the Test Use Policy Co-ordinator/Test Use Committee should be informed about the potential for running an empirical validity study linking the test scores to job performance data. They should decide on whether a validation research study could and should be undertaken in this instance.

1.8 Monitoring Group Differences

- In large scale assessment processes, where sufficient numbers permit, regular reviews of assessment results and their impact on decisions taken with respect to test-takers from different groups (e.g. race, gender, age etc.) should ideally be undertaken. For example, monitoring average group performance differences to investigate test fairness.

2.0 Communication

All test-takers should receive a briefing on the following before completing a psychometric assessment:

- The nature of the assessment;
- Why and how the assessment is being used;
- How long the test-takers' results will be retained for;
- Who will have access to the test-takers' results; and
- If feedback is being offered, when and in what format?

Test-takers should be given ample opportunity to ask questions regarding the assessment process, and receive an honest and professional response.

2.1 Communication with Stakeholders (Non-authorised Users)

- Only authorised test users should decide how test data or test reports are disseminated to stakeholders and/or candidates.
- It is the responsibility of authorised test users to ensure compliance across the organisation with the policies enclosed within this document.
- Where data is to be shared with non-trained stakeholders, authorised users are responsible for ensuring any decision-making and interpretation is made appropriately.

2.2 Informing the Test-Taker

- Test-takers must receive full and relevant information regarding the testing process, prior to completing the assessment. (e.g. confidentiality and use of data).
- Test-takers should never be forced to undertake an assessment, but the consequences of non-participation should be clearly explained (e.g. removal from the recruitment process or development scheme).

2.3 Accommodations - Requesting and Making

- As part of any process where psychometric assessments are used, test-takers should be given the opportunity, prior to completing an assessment, to provide information about any factors or special requirement which could impact on their ability to participate in the assessment process.
- If there is any such requirement, e.g. for a disability, it may be necessary or desirable to make a reasonable adjustment to the assessment process in order to provide the test-taker the same opportunity as others.
- Often the individual themselves is well-placed to advise on suitable accommodations, but we also recommend that the advice of a suitably qualified professional is sought wherever possible. For example, we require dyslexic test-takers to present a certificate from an educational psychologist with a recommendation for additional time.

2.4 Right of Correction

- Test-takers should be given the option to correct any incorrect personal data held on them (e.g. name, title, etc.).

3.0 Test Material & Data Storage

All data stored (on file or electronically) must be held according to applicable legislation.

Non-electronic test data should be stored in locked files, only accessible by authorised test users. Computer-generated test data should be stored in a secure location, only accessible by authorised test users. All personally identifiable data must be stored and transferred in line with CrossTech LLP's data security policy.

3.1 Use of Data

- Where psychometric test data is sought for use in any application other than for the purpose for which it was originally collected, explicit informed consent must be obtained from the test-taker for this additional purpose.

3.2 Confidentiality

- Test data must be treated with the strictest confidentiality.
- Test-takers should be made fully aware of who will have access to their results data, as well as the terms of access (i.e. how their data may be used and for how long it will be stored).

3.3 Authorised Users' Access

- Where the authorised user moves from their current employer or department they should, where possible, ensure that any restricted materials are transferred to another authorised user.
- If there are no qualified users to whom these products can be entrusted you should contact the test publisher immediately to discuss training to become a registered user or, where appropriate, discuss the safe return of materials to the test publisher.

3.4 Copyright

- It is an infringement of ethical guidelines and of international copyright law to copy, modify, translate or in any way replicate, in whole or in part, the products supplied by a test publisher.

3.5 Test Material Storage

- All assessment materials must be stored securely and only be accessible to qualified users.
- An accurate record of all materials held should be maintained to guard against misuse and to ensure the security of all test materials.

4.0 Responsibility for Competence

Authorised users are responsible for maintaining their own level of competence in the use of psychometric assessment tools. Users should not offer services which lie outside their competence. Users are responsible for maintaining their knowledge of current legal and best practice requirements regarding the use of psychometric tools. Users should seek appropriate advice when required, particularly as circumstances begin to challenge their professional expertise. Users should strive to ensure that those working under their direct supervision also comply with each of the requirements of this standard and that these individuals are not required to work beyond the limits of their competence.

4.1 Professional Guidelines

- It is the authorised test user's responsibility to ensure their practice is conducted in line with professional guidelines, including those laid out by the International Test Commission (ITC) and British Psychological Society (BPS).
- Where tests are used in different countries, it is essential to refer to the expertise of authorised users within those countries and to check that practice is in line with appropriate local regulations and guidelines.

4.2 Review of Policy

- Policies should be reviewed as a minimum every two years, or when legislation or other factors dictate the need for a review.

Swift Occupational Ability Pre-Coursework Content with Answers

Module One: Key Figures in Ability Testing

From early in the recorded history of different civilisations, humans have assessed each other.

In the 7th Century, the Chinese developed a method of testing their civil servants that involved written tests assessing verbal ability and creativity.

Samuel Pepys developed navigation tests for the rank of lieutenant in the British Navy in the 17th century, recognising that sailing skills were essential for achieving senior status.

It was not until 1905 that Alfred Binet and Théodore Simon developed one of the first formal psychological ability tests, which was a forerunner of the modern ability tests of today. Whilst researching intellectual disability in French school children in order to identify students in need of alternative education, Binet and Simon developed a test of 30 tasks of increasing difficulty which assessed attention, memory and verbal skill. They developed the concept of 'mental age': the idea that intellectual capability is related to your age as you develop through childhood. This led to the formulation of IQ scores for children.

Since the turn of the 19th Century there have been many theorists who have different views of intelligence, and these represent a few key figures:

Spearman

- In 1905, Charles Spearman observed that those who perform well on one ability test tend to do better on others.
- He proposed a higher factor of general intelligence – which he coined 'g' – complemented by specific abilities. In other words, Spearman viewed intelligence as one overall inherited higher factor.
- Following this thinking – Spearman would argue that someone who scores well on a verbal ability assessment would also score well on numerical and abstract ability assessments. The influence of this thinking is still prevalent in educational psychological testing today.

Thurstone

- In 1924, Thurstone attacked the concept of 'g'.

- He posited a multi-faceted view; that intelligence was composed of seven Primary Mental Abilities.
- These seven abilities were: verbal comprehension, word fluency, number facility, spatial visualisation, associative memory, perceptual speed and reasoning.
- Thurstone's theory places greater emphasis than Spearman's on the fact that people can have high mental ability in one area while being lower in others.

Vernon

- Phillip Vernon is another key name in the history of intelligence and actually reconciled the theories of Spearman and Thurstone, theorising a hierarchical structure of intelligence.
- In 1940, Vernon broke down 'g' into two factors, commonly known as 'academic' and 'practical'.
- The academic factor refers to abilities including reading comprehension and arithmetic reasoning, whereas practical was more focused on mechanical and spatial abilities. These areas of ability can be split into specific skills.
- Interestingly, the English 11+ education system was based on this model of intelligence. If students were to pass their 11+ they would go on to Grammar school, where the focus was academic – and if they failed they would attend a Secondary Modern School where the emphasis was more practical.

Cattell

- After Vernon reconciled previous theories, three years later, Cattell diverged once more. Up to the 1940s, intelligence was thought to be inherited, until Raymond Cattell came up with the concepts of Fluid and Crystallised Intelligence. Psychologists have since argued about the nature vs. nurture debate for many years.
- Fluid Intelligence was defined as the ability to deal with novel and abstract problems. It was thought to be genetic and therefore immune to culture and environment (e.g. it could not be taught), and thought to decline with age. Abstract reasoning would therefore be a cognitive ability test aligned to this type of intelligence.

- Crystallised Intelligence, however, was defined as intelligence grounded in knowledge, expertise and wisdom learnt over time, and thus was thought to increase with age. Saville Assessment's aptitude tests do not test knowledge explicitly; prior knowledge in specific areas is not required to complete a verbal reasoning test for example.
- The concept of Crystallised Intelligence is important when considering occupational learning: through increasing your exposure to working with information and by learning new knowledge or skills, this can help you to develop your abilities in certain areas; for example, gaining more exposure to working with numerical data over time could help an individual improve their ability in this area.
- Ironically, tests that have been developed to be culture fair actually turned out to be biased towards white middle-class individuals.

Gardner

- In 1983, Gardner proposed his theory of Multiple Intelligences, where some intelligences are best measured using specific skills assessments; for example, music tests to assess musical intelligence, or personality tests for inter-personal and intra-personal intelligence.
 - 'Linguistic' encompasses words and their meanings
 - 'Logical' focuses on reasoning and drawing conclusions
 - 'Spatial' refers to visualising shapes in space
 - 'Musical' intelligence covers musical composition and performance
 - 'Kinaesthetic' covers co-ordinating bodily movements
 - 'Inter-personal' intelligence focuses on the need to understand and get on with people
 - 'Intra-personal' intelligence covers the ability to control one's inner self

So, as we have seen, there have been many shifts between singular and multiple conceptualisations of intelligence. In light of this, we will now take a look at where this debate sits today.

General vs Specific Intelligences

To this day, it is still debated whether intelligence is a single construct or formed of multiple constructs. Some academics favour a singular form of intelligence. As discussed, this would mean that someone who is good at one subject is likely to be good at another. For example, if someone can play the piano well, they are also likely to be able to pick up playing the violin easily too.

An example of a test that could be used to assess general, or singular, intelligence would be an IQ Test; the Wechsler Adult Intelligence Scale, for example. This test measures intelligence over four broad areas: Verbal Comprehension, Perceptual Reasoning, Working Memory and Perceptual Speed, and combines these to give a Full Scale IQ score. The fourth edition of this test, which was developed in 2008, is often used by neuropsychologists to assess overall cognitive function of the brain.

Some applied psychologists and practitioners lean towards a multi-faceted view of intelligence. An example of tests used to measure specific ability areas would be aptitude tests; a verbal reasoning assessment or an error checking test, for example. Practitioners tend to have a concern for the relevance of a test to the job, rather than being concerned with theoretical considerations.

Whilst evidence shows that tests of general intelligence predict job performance just as well as tests of specific abilities, in an occupational setting, it makes practical sense that you are only measuring abilities which are found to be relevant to the role. It is important to understand the level of general ability and the kind of specific abilities that are required for a role; assessing the related attributes of an individual enables sound judgements about job-fit to be made, and measuring specific abilities that have clear job relevance results in greater candidate, manager and legal acceptability.

On this basis, Vernon's hierarchical model integrating 'g' with more specific abilities - under the umbrellas of practical and academic intelligence - carries great merit.

Considering the applications of these theories of intelligence, cognitive ability tests are generally found to be the strongest single predictors of work performance across job roles, compared to other selection methods. This lends them well to the applications of workplace selection, development and career guidance.

Kuncel et al. (2010) found that cognitive ability predicts academic performance and, perhaps more surprisingly, broad life outcomes such as occupational attainment and stable employment, and even mortality, in addition to job performance.

Saville Assessment do not focus on extremely broad ranging intelligence measures, such as the Wechsler Adult Intelligence Scale measure of IQ. We do, however, offer both measures of single aptitude and combined aptitudes (Swift tests) which assess multiple types of aptitude through sub-tests in one assessment sitting. The single aptitude tests allow you to target an in-depth cognitive aptitude that

is central to success in a role, e.g. a numerical aptitude test for an accountant. However, if there are several important aptitudes related to success, single tests can be a more time-consuming method to assess these, and combined tests such as Swift target the relevant areas more broadly and efficiently. Single tests give a score for the single aptitude area which can be used for decision making. The decision-making score from a combined Swift test is the overall score (rather than scores from sub-tests).

Next, we'll look at some of the different ways we can measure ability or intelligence.

Notes:

Questions and Answers

Q1 Some argue intelligence is a single construct. Others believe it is formed of multiple constructs. Pull the correct answer into the relevant box to match each of the approaches with a brief explanation and an example of a relevant theory.

Approach to Intelligence	Explanation	Relevant theory
<i>Single construct</i>	People who are strong in one ability are also strong in others.	Spearman's higher factor of general intelligence (g)
<i>Multiple construct</i>	People may be strong only in one specific ability area (e.g. have strong numerical aptitude and weak verbal aptitude)	Thurstone's seven Primary Mental Abilities

Q2 Which of the following is true of using tests that measure specific abilities in an occupational setting?

You can avoid assessing areas of ability/intelligence which are not relevant to the role.

Q3 Cattell proposed two different types of intelligence; fluid and crystallised. Match the following descriptions to the type of intelligence.

Type of Intelligence	Features of the Theory
<i>Fluid intelligence</i>	The ability to deal with novel and abstract problems. Largely genetic and culture-free. Declines with age.
<i>Crystallised intelligence</i>	Intelligence grounded in knowledge, expertise and wisdom. Learnt over time. Increases with age.

Q4 What has cognitive ability been found to predict?

Work performance, educational performance and occupational attainment (all of the above)

Q5 Ability tests would be appropriate for use in which of the following applications? Tick all that apply.

Career guidance

Workplace development

Recruitment

Q6 Which of the following is an example of a single test of ability?

Numerical aptitude test, e.g. Saville's Numerical Comprehension test

Module Two: Types of Assessment

In this module we will consider two different types of assessment: those which are concerned with the performance of others, typically known as norm-referenced tests, and those which relate to an external standard, sometimes known as domain- or criterion-referenced tests.

Externally-referenced Tests

Externally-referenced tests ask “How well does the candidate compare against an expected standard?” Candidates have to reach a required standard on the test, irrespective of how anyone else performs. A good example of this is a driving test: to pass, one must meet a set required standard on the test to be given a driving licence. This type of testing is also common in education – students generally are expected to reach a set standard to achieve a certain qualification or grade.

Work samples are one example of a type of test that can be externally referenced. Externally-referenced work samples require candidates to perform a task that is similar to or the same as the job with candidates expected to meet a minimum standard of performance. For example, an individual completing a typing test may have to be able to type at least 35 words per minute to pass. Work samples can be particularly useful where on-the-job training is less of an option.

Work samples are good predictors of workplace performance. They are a useful inclusion in some recruitment processes because they simulate behaviour or skills required on the job. Some work samples focus on an individual’s potential to learn tasks during the test itself; these are often known as trainability tests.

Because work samples are costly to develop they are particularly well suited to volume recruitment. For some roles it is straightforward to measure critical job tasks that form a large proportion of the role. For example, checking for faults on a manufacturing assembly line is relatively easy to simulate. Work samples are also useful where skill mastery is critical, for example using flight simulators in the selection of pilots. While useful, work samples are not appropriate for all roles, for example they do not adequately capture many of the key tasks required in a typical management role.

Norm-referenced Tests

Norm-referenced tests compare an individual’s performance against the performance of others. Aptitude tests are one type of ability test which are typically norm-referenced and indicate an individual’s level of aptitude

compared to other people. Aptitude tests that use norm referencing focus on predicting the future performance of individuals and typically do not require the test-taker to have specific knowledge or experience to do well. For example, if an individual applying to a graduate scheme were to take a numerical reasoning aptitude assessment the results would benchmark them against other graduates who have taken the test. This type of referencing is popular in recruitment where selection is competitive and must only be done on a relevant task. You will learn a lot more about norm-referencing on the course.

At Saville Assessment we have a wide range of ability tests, each of which is suitable for particular industries and levels of an organisation. Married with this, we have a variety of comparison groups which can be chosen to best match the applicant pool. Choosing the appropriate assessment and comparison starts with a thorough analysis of the job role. Doing so will not only ensure you are measuring the specific competencies and skills which are relevant and predictive of job performance, but the tests are likely to feel more relevant to candidates.

Saville Assessment’s approach will be discussed in more detail in Module Three.

Questions and Answers

Q7 Which of the following is an example of an externally-referenced test?

An exam where the individual is required to meet a set standard to pass

Q8 Which of the following is an example of a norm-referenced test?

A mechanical reasoning aptitude test showing scores in comparison to others

Q9 When might you develop a work sample test? Tick all that apply.

When you have a large volume of applicants for a role

When you can simulate a large important part of the job

When there is a need for mastery of a particular skillset

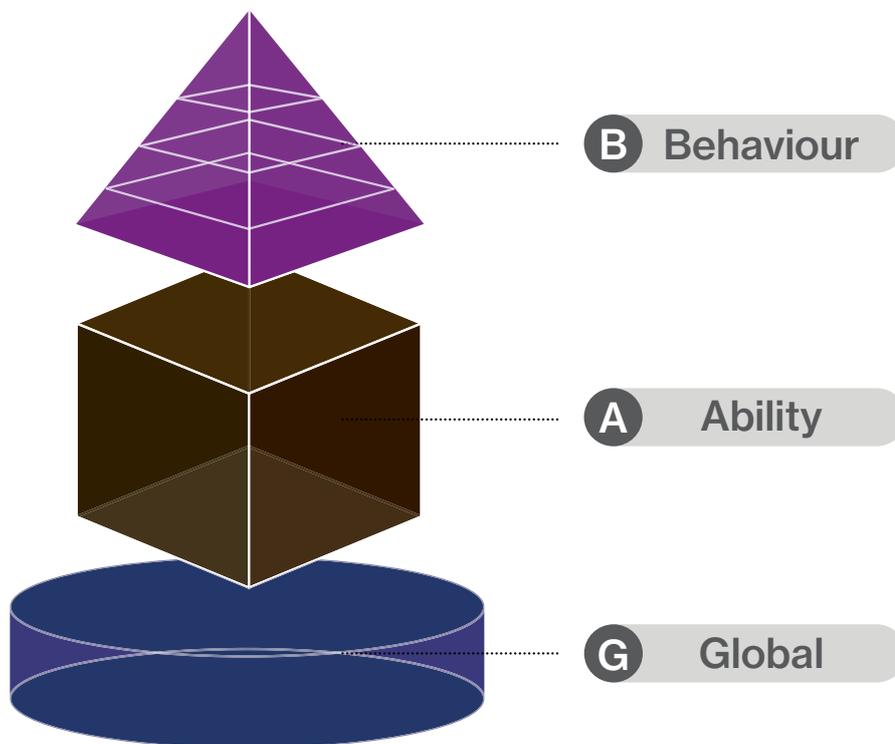
Module Three: The Saville Assessment Approach

Our Product Portfolio

Now that we've covered the main theories of intelligence and types of assessment, we will explain our approach, including the model that underpins our aptitude assessments.

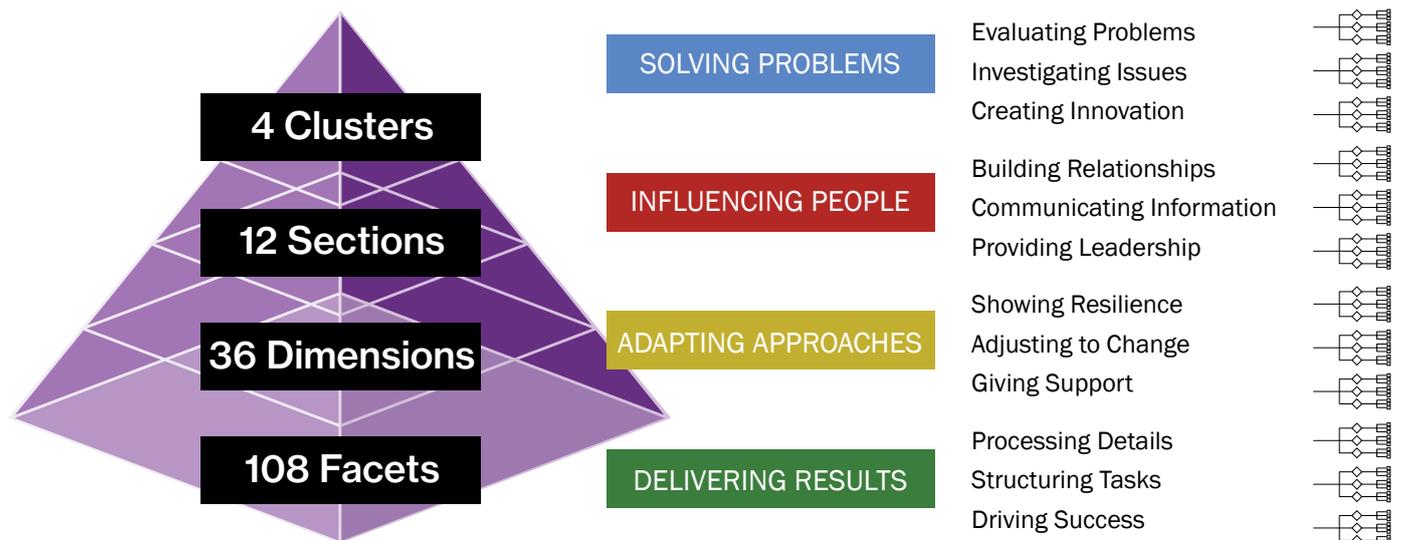
Saville Assessment B-A-G Framework

To better conceptualise performance at work, we break our framework down into three separate aspects: Behaviour, Ability and Global.



Behaviour

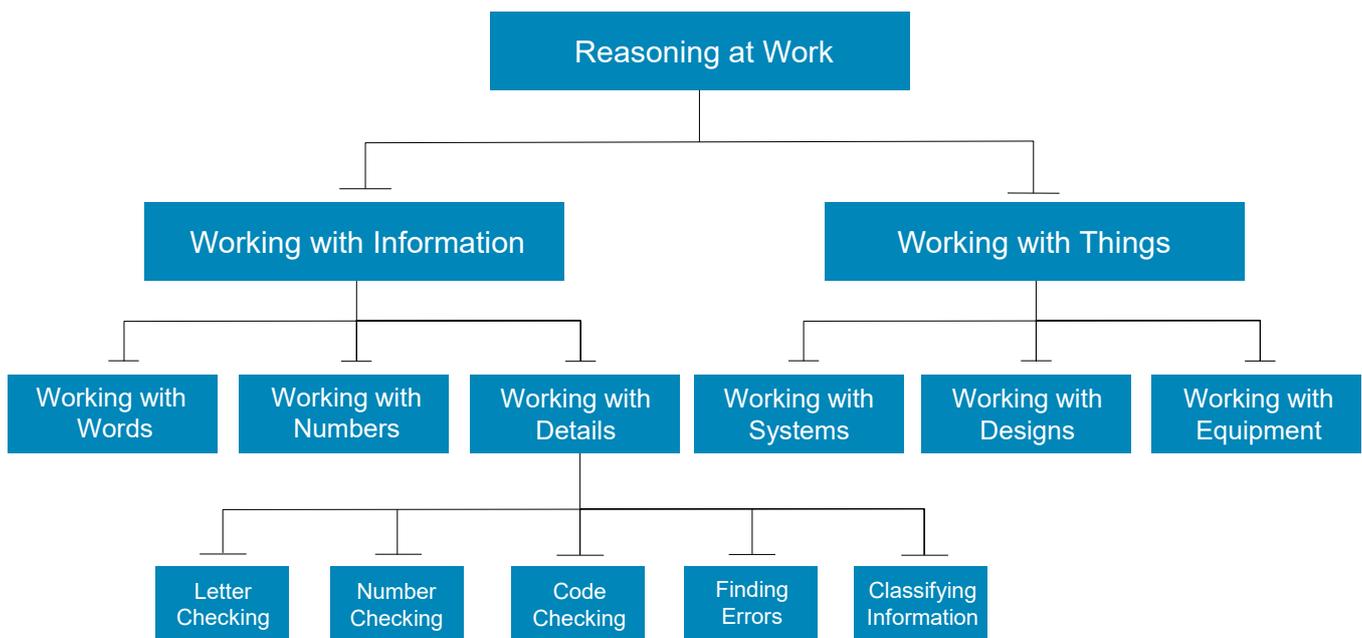
We primarily measure behaviour using our Wave assessments. The behavioural element of our B-A-G model is hierarchical, comprising four broad clusters of workplace behaviour that can be drilled down into more specific elements. The four clusters are split into 12 sections that comprise 36 dimensions, which are underpinned by 108 specific facets of workplace behaviour. These behaviours were selected as the strongest predictors of effective workplace performance. This will all be covered in much more detail for those of you who will be undertaking our Wave course.



Ability

Our model of Ability is hierarchical in nature and represents the abilities required at work, and is closely linked to the underlying aptitudes measured through the tests in the Saville Assessment Aptitude Assessment portfolio. The model is made up of one overall cluster – Reasoning at Work (comparable to Spearman's g). This breaks down into two sections: Working with Information and Working with Things (mirroring Vernon's Academic and Practical factors). This forms the scientific basis for our aptitude portfolio.

Underpinning Working with Information and Working with Things are six ability dimensions, which are measured through our aptitude tests. The six dimensions are further defined into specific areas of workplace ability. For example, Working with Details is measured by our Error Checking assessments which are comprised of questions covering Letter Checking, Number Checking, Code Checking, Finding Errors and Classifying Information.

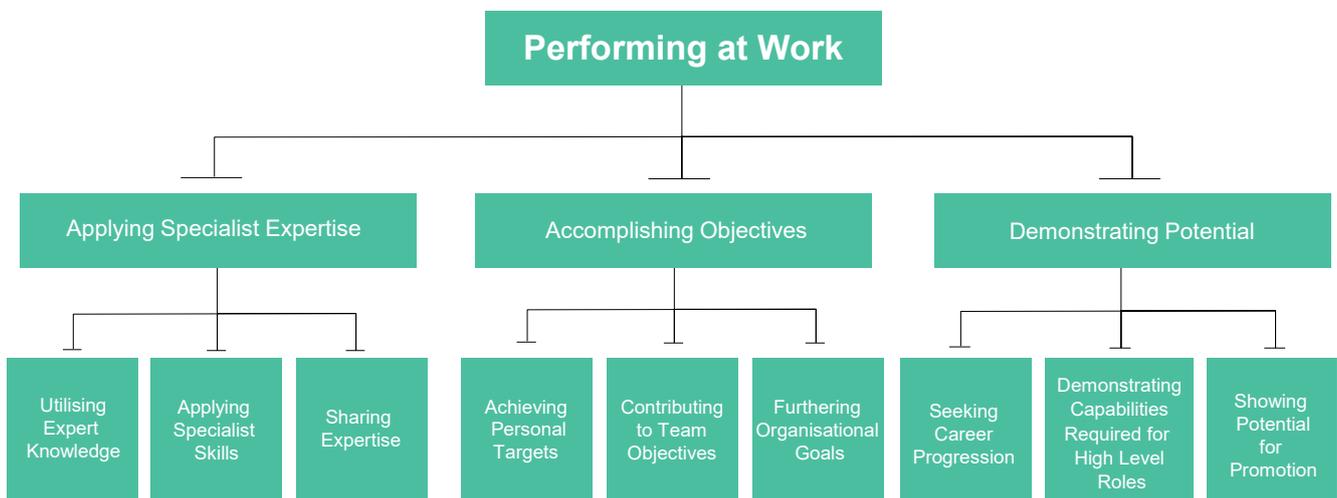


Global

'Global' describes broad overall effectiveness characteristics of performance and potential at work.

Our Global aspects of performance at work define performance at a broad, high level across roles. The focus is on overall effectiveness in jobs, rather than narrow criteria which can vary from job to job.

Overall effectiveness at work comprises Applying Specialist Expertise, Accomplishing Objectives and Demonstrating Potential. Applying Specialist Expertise refers to technical expertise and Accomplishing Objectives refers to a drive to achieve targets and objectives. Demonstrating Potential relates to the capacity of an individual to be effective in the future in terms of being able to take on higher level or broader roles.



The Talent Cycle

The B-A-G model discussed underpins all of Saville Assessment's assessment solutions across the talent cycle. This integrated model of assessment uses the same language throughout the talent cycle to measure an individual's performance and potential in critical work areas.

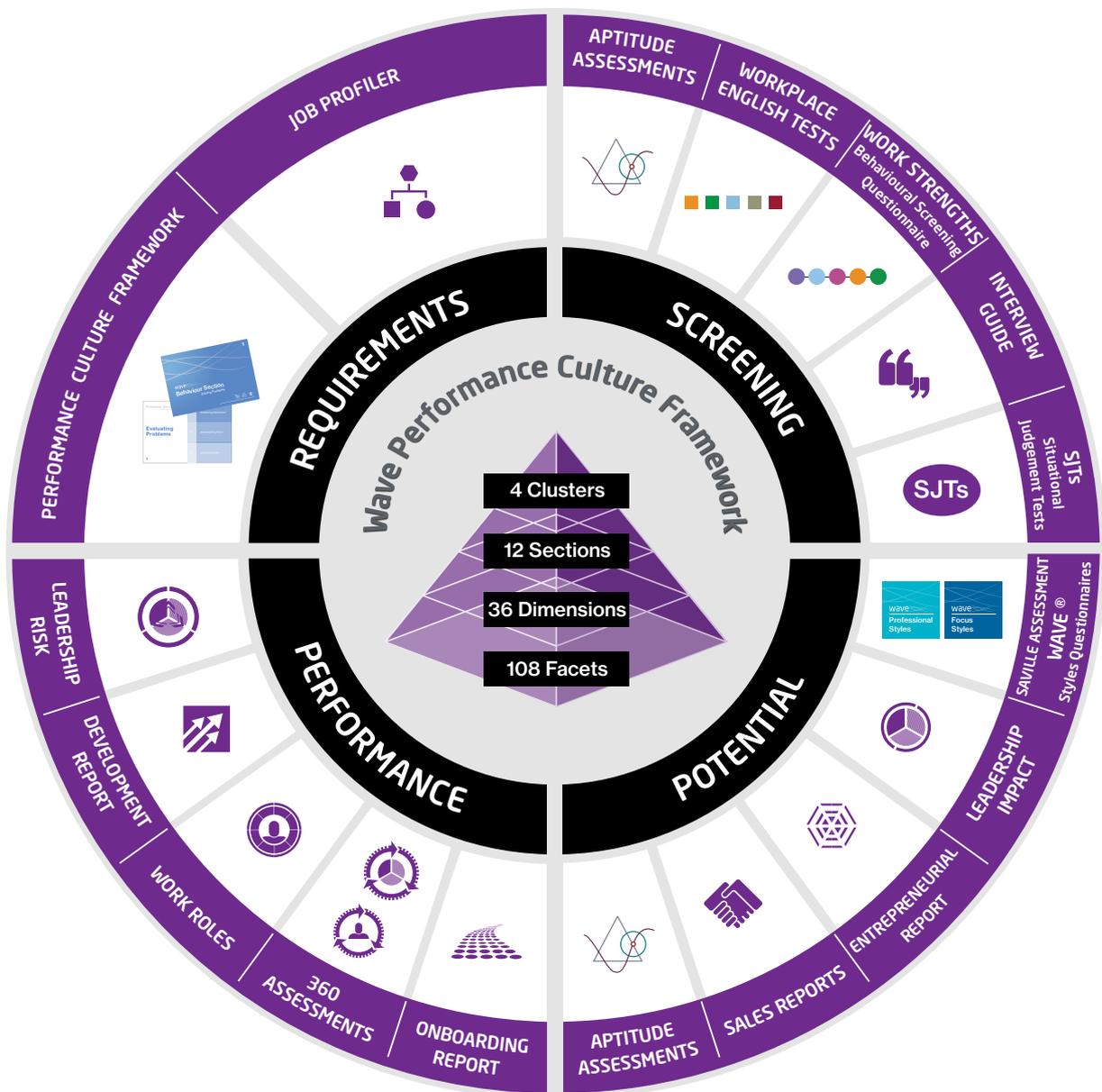
The Saville Assessment talent cycle comprises four quadrants: the first area is profiling requirements. Before embarking on any assessment process, you need a clear understanding of what you are looking for. Saville Assessment's profiling tools help users to engage relevant stakeholders in defining the key criteria for success in a role.

The next quadrant focuses on screening, where we seek to identify the candidates who are most suitable for particular roles. Saville Assessment's aptitude assessments,

situational judgement tests and Work Strengths behavioural screening questionnaire support candidate screening, increasing efficiency in volume recruitment whilst creating an engaging candidate experience.

The area of potential covers a number of applications such as selection, identifying high potential staff, and developing staff. The Wave Styles questionnaires provide an understanding of motives and talents in critical work areas and there is a wide variety of reports that can be generated to support different applications. Aptitude assessments also offer a highly predictive indicator of future work potential.

Lastly, the performance quadrant contains different reports, powered by our Wave Styles and Wave 360 assessments, that help individuals or teams progress and develop. The Wave Styles assessments will be covered in more detail on the Wave course.



Background to Saville Assessment

The journey of Saville Assessment started in 2004 when a team of assessment specialists came together. The team comprised experts in Occupational Psychology, Business Assessment and Information Technology, with the goal of transforming and revolutionising assessment around the world. So what does this mean in practice?

Transforming Assessment Around the World – Saville Assessment advanced traditional assessment approaches of the 20th century by creating innovative, rigorous and work-relevant assessment tools, including online ability and personality measures and multimedia simulations.

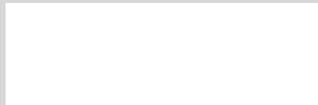
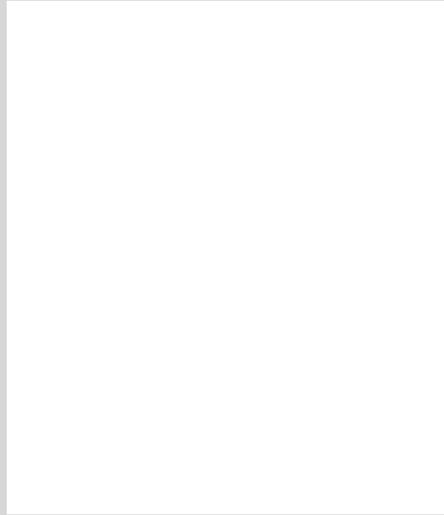
Maximising the Power of the Internet – Instead of having questionnaires and assessments that were designed for paper-and-pencil use, and placing them on the internet, our tools are designed from first principles to reap the benefits of online assessment.

Greater Insight and Richer Feedback – Our assessments are designed not only to be more accurate and detailed than any other, but also facilitate clear and valid interpretation, giving further clarity and detail into every individual who is assessed.

Taking Seriously the Need for Security in Testing - We place security at the core of our assessments, exemplified by our online aptitude tests which are designed to counter cheating.

Tailoring our Approaches to Fit Organisations' Needs – Many of our tools can be transposed on to a target organisation's own model of performance, harnessing our research-based measurement power and delivering in an organisation's own language.

Exceptional Prediction of Work Performance – Our approach is driven by data to predict a clear model of performance at work. So, when you use a Saville Assessment psychometric test, you can make strong and valid predictions about work performance.



About Saville Assessment, a Willis Towers Watson Company

Our mission is to transform assessment around the world. We enable organisations to identify potential, accelerate performance, and achieve outstanding results. Our portfolio of leading-edge assessments are designed based on extensive research into successful workplace performance and the critical relationship between motive, talent and workplace culture. With representatives in over 80 countries we are transforming how organisations Hire, Build and Lead talent globally. Learn more at www.savilleassessment.com