

2.8 Intelligence and Achievement

Name: _____ Class: _____ Date: _____ **Total: 18 marks**

KEY WORDS intelligence 智力 • reliability 信度 • normal distribution 正态分布
• stereotype threat 刻板印象威胁 • validity 效度

Objective

Build the skills to answer exam questions on **intelligence** 智力 and its measurement.

You must be able to:

- state what **general intelligence** 一般智力 claims and one alternative view
- explain **reliability** 信度, **validity** 效度 and **standardisation** 标准化
- describe a **normal distribution** 正态分布 of scores
- discuss **stereotype threat** 刻板印象威胁 and test bias

1 Worked examples

Study these first. Each one shows the method for a task used later.

■ One factor or several

The **general intelligence** view holds that performance across different mental tasks correlates, implying a common factor. Alternative theories propose several relatively independent intelligences or types, and disagree about how much a single score can capture.

■ Three test properties

Reliability: the test gives consistent results. **Validity:** it measures what it claims to measure. **Standardisation:** scores are interpreted against a representative comparison group. A test can be reliable and still invalid.

■ The distribution

Standardised intelligence scores form a **normal distribution**: most people cluster near the average, with symmetrical tails. The mean and standard deviation are set by the standardisation sample.

■ Stereotype threat

When a negative stereotype about one's group is made salient, performance can drop. This is a situational effect on the **score**, not evidence about ability, and it is one reason score gaps require careful interpretation.

2 Practice

2.1 State the difference between reliability and validity. [2]

2.2 Explain what standardisation of a test involves. [2]

2.3 A test gives the same score to the same person on two occasions, but the score does not predict school performance at all.

(a) State which property it has. [1]

(b) State which it lacks. [1]

3 Exam-style questions

3.1 In a normal distribution of test scores, most people score [1]

- **A** near the extremes
- **B** near the mean
- **C** identically
- **D** above the mean

3.2 Reminding candidates of a negative stereotype before a test can lower their scores. This is [1]

- **A** evidence of lower ability
- **B** stereotype threat
- **C** test-retest unreliability
- **D** standardisation

3.3 A test is standardised on one national group and used with another.

(a) State one problem this may create. [1]

(b) Explain why a score difference between the groups would be hard to interpret. [2]

3.4 Applied research task. A test of reasoning is given to 2 000 people. Scores form a normal distribution with a mean of 100 and a standard deviation of 15. The test is re-administered to the same people six months later; the two sets of scores correlate at $r = 0.91$. Scores on the test also correlate at $r = 0.62$ with school grades three years later.

(a) State what percentage of people score between 85 and 115, and explain your answer. [2]

(b) Name the property shown by the $r = 0.91$ result and the property shown by the $r = 0.62$ result. [2]

(c) Explain the difference between fluid and crystallised intelligence, and state which the test is more likely measuring. [2]

(d) Explain what stereotype threat is, and how it could lower a group's average score on this test without any difference in ability. [1]

Solutions

2.1 reliability is consistency of results; validity is whether the test measures what it claims to.

2.2 administering the test to a representative sample so that later scores can be interpreted against that comparison group.

2.3 (a) reliability. (b) validity.

3.1 B. a normal distribution clusters around the mean with symmetrical tails.

3.2 B. the situation, not the ability, has changed.

3.3 (a) the comparison group may not represent the people being tested. (b) the difference could reflect test content, language, familiarity or stereotype threat rather than ability, so the score alone cannot identify a cause.

3.4 (a) About **68%** (1); 85 and 115 are one standard deviation below and above the mean, and in a normal distribution roughly 68% of scores lie within one standard deviation of the mean (1).

(b) $r = 0.91$ across two administrations shows **test–retest reliability** — consistency of measurement (1). $r = 0.62$ with a later outcome shows **predictive validity** — the test predicts what it claims to (1).

(c) **Fluid** intelligence is the ability to reason about novel problems, independent of learned knowledge, and it declines with age (1); **crystallised** intelligence is accumulated knowledge and skill, and it holds or increases with age. A reasoning test is more likely measuring **fluid** intelligence (1).

(d) Stereotype threat is the anxiety of confirming a negative stereotype about one's own group (1); that anxiety consumes working-memory capacity during the test, so performance falls even though underlying ability is unchanged.