

**WRAP Exams · wrapexams.com**

Prelims PYQ, Mains model answers, and copy evaluation. Write to the desk for this paper, a doubt, or the AI waitlist.

[ask@wrapexams.com](mailto:ask@wrapexams.com) · +91 85951 84903 · [wrapexams.com](https://wrapexams.com)

Q4(c)

## **What is reliability ? Explain the different tests available to social science researcher to establish reliability.**

UPSC Civil Services Mains 2022 · Sociology GS 1 · 15 marks · Works and Economic Life

Works and Economic Life: Social organization of work in different types of society; Formal and informal organization of work; Labour and society.

### **Core demand of the question**

- What is reliability? Explain the different tests available to social science researcher to establish reliability
- Keep the explanation within a 20-mark length

### **Revision summary**

Reliability is consistency, not truth.

**Test-retest**, parallel forms, split-half, Cronbach's alpha, and inter-rater agreement are the main tests.

**Durkheim's** suicide files already needed stable coding.

**Census** definition changes can mimic social change.

A reliable scale can still miss **Weber's** intended meaning.

### **Introduction**

Reliability is the consistency of a measure. If the same instrument, applied to the same kind of object under the same conditions, jumps about, the scores are unreliable. Validity is a later question: a clock can tick reliably and still tell the wrong time.

### **Body**

#### **What reliability is**

- In survey and test work, reliability is the share of variance that is not random error.
- **Emile Durkheim** already needed stable classification of deaths before suicide rates could be compared. Unstable coding is an reliability failure in official data.
- **Qualitative work uses a cousin idea:** inter-coder agreement and stable use of a codebook, not a claim that meanings never shift.

#### **Test-retest**

- The same test is given twice to the same persons after an interval.
- Use when the trait should be stable, such as a literacy item, not a mood item after a riot.
- Memory and real change both threaten the correlation.

#### **Parallel or alternate forms**

- Two versions of the scale are built to the same blueprint and both are given.

- Use in examinations and long batteries to reduce item memory.
- Costly to write well.

### Split-half and internal consistency

- The test is split, scores are correlated, and Spearman-Brown may correct the half-length.
- Cronbach's alpha estimates consistency among items of a Likert battery.
- Use for attitude scales on, say, gender-role items in a student sample. Alpha is not a test of unidimensional truth, only of hanging together.

### Inter-rater or inter-coder reliability

- Two observers code the same focus-group transcript or **Census** occupation string.
- Cohen's kappa or percent agreement is reported.
- Essential for content analysis, courtroom files, and open-ended NFHS-type comments.

### Conditions and cautions

- High reliability does not prove that "caste" or "work" was well conceptualised. That is validity, and **Weber's** meaning problem.
- Official series such as the **Census** need stable definitions across decades; a definition change looks like social change if reliability of the series is ignored.
- Hidden populations reached by snowballing cannot borrow NFHS reliability claims without a design argument.

### Flow diagram

Reliability consistency -> Test-retest  
Reliability consistency -> Parallel forms  
Reliability consistency -> Split-half alpha  
Reliability consistency -> Inter-rater  
Reliability consistency -> Validity is separate

### Conclusion

Reliability is consistency of a measure across time, forms, items, or raters. **Test-retest**, parallel forms, split-half, alpha, and inter-rater tests are the standard kit. Use them, then still ask whether the consistent score measures the sociological object.

### Facts & figures

#### Test-retest

Correlation of two administrations; threatened by memory and by true change.

#### Cronbach's alpha

Internal-consistency estimate for multi-item scales; high alpha is not proof of validity.

#### Inter-rater reliability

Agreement of independent coders on the same material; central for qualitative and official coding.

**Census series**

Comparability of work and household definitions is a reliability problem for trend claims.

---

WRAP Exams · Write. Read. Analyse. Practice. · <https://wrapexams.com/upsc/mains-answers/sociology-gs-1/2022/what-is-reliability-explain-the-different-tests-available-to-social-science-researcher-to-establish-reliability/> · [ask@wrapexams.com](mailto:ask@wrapexams.com) · [wrapexams.com](https://wrapexams.com)

WRAP Exams