

Knowledge organiser:

Reliability

What is reliability?

Internal reliability: the extent to which a test or measure is **consistent within itself** e.g. the use of standardised instructions and procedures for all participants.

External reliability: the extent to which a test produces **consistent results over several occasions**.

Assessment

AO1: Demonstrate knowledge of internal and external reliability, ways of dealing with reliability issues and assessing reliability.

AO2: Apply knowledge of reliability, its issues and ways reliability can be assessed, to novel scenarios, developmental (Kohlberg) and social (Milgram) psychology, and to the personal investigations (Unit 4 only).

What are reliability issues?

Lack of operationalisation of the variables is a potential issue. If you do not clearly state how you are defining and measuring your variables, this makes replication impossible and makes inter-rater reliability issues more likely (see below for details). *Order effects* can occur in a repeated measures design. Whilst completing the second condition the participant can become practiced/better at the task the second time round, or they can underperform due to factors such as boredom. A lack of *standardised procedures, inconsistency of measuring tools and an uncontrolled environment* all jeopardise reliability.

How do we deal with reliability issues?

Increasing levels of control in a study raises reliability – using a *standardised procedure, instructions and scientific measuring equipment* (e.g. PET/ MRI scans over self-reported ‘feelings’) and conducting the research in a *laboratory environment* helps to overcome most reliability issues. The greater the level of control, the lower the risk of extraneous variables effecting results.

Order effects can be combatted with *counterbalancing*. This involves splitting the population/sample into two halves; the first half complete condition one followed by condition two, the second half complete condition two followed by condition one. This in effect negates the impact of order/practice effects, as both conditions are affected equally, and increases internal reliability.

Assessing reliability: measures used to check the consistency of a set of results/findings.

Split-half reliability/testing: involves splitting a participant’s test answers in half and seeing whether s/he got the same or similar scores on the two halves. If so, internal reliability is high; if not, it is low and individual questions would need to be re-ordered or redesigned to improve reliability.

Test-retest reliability: involves testing and retesting the same participants over time, with the same test, and comparing their scores. If the scores are the same, the test has external reliability.

Inter-rater reliability: where two or more psychologists produce consistent results by using a standardised procedure, agreed coding system, or correlation of their data.