Unpacking the terminology – Data and Measurement in VCE Psychology

Drawing evidence-based conclusions and evaluating claims involves the analysis of data and identification of sources of uncertainty and possible bias.

Measurement terms related to the analysis and evaluation of quantitative data are defined on pages 19 and 21 of the *VCE Psychology 2023–2027 Study Design*. Students are expected to apply measurement terms to the analysis, interpretation and evaluation of their own and others’ investigation data. Further advice and examples are provided in this document.

The relationship between accuracy and true value

Accuracy refers to the closeness of a measured value to the true value (or standard value/widely accepted value where a ‘knowable’ true value is not possible). Accuracy can therefore be considered from the following perspective: if a quantity could be measured with perfect accuracy, then we would find its true value. The nature of human behaviour and mental processes measured in the discipline of psychology also often involve psychological constructs where a true or accepted value may be difficult to quantify. For example, in an experiment investigating stress levels we may not be able to define an ‘accurate’ measure of stress, instead we may measure something we associate with stress like the heart rate of the participants. This may be measured using a heart rate monitor. The true value of their heart rate would be the value obtained if it was measured perfectly, say 85 bpm. If the heart rate monitor detected a heart rate of 75 bpm instead, this would be inaccurate. If the true or accepted value for a measurement is not known, comments on accuracy would not be valid.

The relationship between accuracy and precision

Measurement results that are close to the true, or accepted, value of a quantity are said to be accurate. A set of measurement results that have very similar or the same values are said to be precise. Following on from the previous example, if a participant’s heart rate was to be measured every five minutes, in unchanging conditions, you would expect there to be little fluctuation in their heart rate. If their heart rate was recorded as 85 bpm every five minutes over the course of an hour, then the experimental data was precise and accurate. However, if there is no indication of how close the results are to the true value or if there is no standard reference against which these measurement results can be compared, then the degree of accuracy cannot be determined. The degree of precision can always be determined, provided multiple measurement results are recorded.

A set of measurement results may therefore be:

* both accurate and precise
* accurate but not precise
* not accurate but precise
* neither accurate nor precise.

Both accuracy and precision are qualitative terms: when comparing measurement values, they can be said to be ‘more’ or ‘less’ precise or accurate than each other.

A common analogy used to [illustrate](https://www.researchgate.net/figure/Precision-versus-accuracy-The-bullseye-represents-the-true-value-eg-the-true_fig6_304674901) the difference between accuracy and precision is to imagine throwing darts at a dartboard, with the centre of the dartboard representing the ‘true value’. Darts that are thrown close to the bullseye are said to be accurate, while darts that are grouped together are said to be precise.

The relationship between repeatability and reproducibility

Both repeatability and reproducibility can be used to evaluate the quality of data in terms of the precision of measurement results, although reproducibility is a more difficult test of the quality of data. That is, it is more likely an experiment will be repeatable than it will be reproducible.

Repeatability refers to the degree of agreement obtained when a set of replicate measurement results are produced by the same student (or group of students working together as a team), using the same method and equipment, under the same conditions, over a short timeframe. A measurement can be described as ‘repeatable’ in quality when repetition under the same conditions gives the same or similar results. In general, scientists perform the same scientific investigation several times to confirm their findings; this practice is also important for VCE Psychology students wherever practicable.

Reproducibility refers to the degree of agreement between the results of investigations conducted by different students (or groups of students working together in teams), working under different conditions, generally using different equipment at different times and often using different methods of investigation. A measurement can be described as ’reproducible’ in quality when a new investigation under equivalent (but not identical) conditions produces the same or similar results. Reproducibility is particularly important for scientists because it can confirm the findings achieved in one laboratory or location by other scientists elsewhere in the world and establishes methods that other scientists can modify and enhance (this in turn relates to external validity).

Reliability

The *VCE Psychology 2023–2027 Study Design* does not refer to the ‘reliability’ of quantitative data because of the ambiguity of its use, both in everyday usage and in science texts where it is used to refer to raw data, data patterns and conclusions, as well as to information sources. The term ‘reliability’ is not found in the standard VIM or GUM metrology references. For quantitative data, the terms ‘repeatable’ and ‘reproducible’ are clearer and therefore more useful for students to understand and apply in analysing and evaluating their own investigations.

Personal errors

Personal errors include mistakes or miscalculations, such as students misreading a scale on a thermometer as 35 oC rather than 25 oC, or height as 146cm rather than 148cm. Repeating measurements is good scientific practice that allows students to detect and correct mistakes and miscalculations.

Random and systematic errors

The effects of random and systematic errors contribute to uncertainty in measurement. Random errors affect precision. Systematic errors (both unknown and unadjusted known errors) affect accuracy.

An effect is ‘random’ if it is equally likely to result in the measured value being above or below what would usually be measured. If a participant was required to complete a memory test once a week for a month to test the effect of a cognitive enhancing drug, that participant may perform poorly on each subsequent test due to boredom. This would be a random error, as for each measurement this impact could result in their scores being above or below their true value (actual memory ability). The effects of random errors can be reduced by taking actions such as making multiple measurements and calculating a mean or increasing sample sizes.

An effect is ‘systematic’ if it biases the result (i.e. takes it in one direction) above or below what would usually be measured. Causes of systematic error include measuring instruments that are incorrectly calibrated or used incorrectly, and a flawed experimental method. For example, using an EEG that has been incorrectly calibrated to measure a participant’s brain waves may result in frequency and amplitude readings that are consistently too high or too low by a specific amount. It is not always the case, however, that systematic errors are always a consistent magnitude away from the true value, depending on the cause of the error. For example, in terms of a flawed experimental method, if participants consuming a probiotic report lowered anxiety levels, it may be due to the expectation that the probiotics will have a beneficial impact on anxiety, rather than the biological mechanism of the probiotics. This would be a systematic error, as it biases the result above what would usually be measured (greater change in anxiety levels). The effects of systematic errors can be reduced by taking actions such as making sure that measuring instruments are operating correctly and checking experimental method to ensure that an incorrect procedure or technique is not being repeated.

Students should be able to distinguish between random and systematic errors, and to explain the effects on the quality of the data (precision and accuracy) of these errors in their own and others’ experiments. For example, if blood alcohol concentration measurements are taken with a breathalyser that has been incorrectly calibrated, students should be able to identify whether this would cause random or systematic error effects, determine in which direction this error would affect their final measurement results and then use this determination to predict whether their final results will be slightly higher or lower than what they were expecting. As a learning activity, students could design an investigation to test their predictions and to suggest further investigations that may be required to confirm their ideas or extend their understanding of the effects of errors on measurement data. For those students whose measurements did not match their expectations, they should try and speculate why, suggesting further investigations to confirm their ideas.

Uncertainty

All scientific efforts are directed towards reducing the degree of uncertainty in the world about observations, relationships and causes. VCE Psychology requires only a qualitative treatment of uncertainty (no calculations or statistical analysis).

All measurements are subject to uncertainty and may have many potential sources of variation, including human capacity in measurement and the nature of the instrumentation used, and this uncertainly extends to all inferences and conclusions that depend on uncertain measurements. This is particularly prevalent in the study of Psychology due to both the subjective nature of many of the measurements used as well as the fact that much of what we study is a psychological construct that is believed to exist but cannot be directly observed.

When evaluating personally sourced or provided data, students should be able to identify contradictory (incorrect data) and incomplete data (missing data – questions without answers or variables without observations), including possible sources of bias.

Outliers

Outliers in data are data points that do not fit a pattern or trend. Outliers should not be ignored or assumed to be personal errors (mistakes), but instead should be considered as important observations to account for. All data, including outliers, should be included in order that the analysis of data is undertaken in an ethical and systematic way. It is good scientific practice to take repeated readings so that greater confidence in the validity of data is achieved. More observations will minimise the effect of outliers caused by random factors (such as personal errors in measurement).