Unpacking the terminology –   
Data and measurement in   
VCE Environmental Science

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 16 and 17 of the *VCE Environmental Science 2022–2026 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

If a quantity could be measured perfectly, then we would find its true value. The true value of a quantity is unknown but is what one tries to measure. Accuracy refers to the closeness of a measured value to the true value. Accuracy also refers to how close a measured value is to a standard value or a widely accepted value. For example, given an accepted value for the heat of combustion of methane gas at 25 oC and 100 kPa pressure of 55.6 kJ g-1, if Student A experimentally determines the heat of combustion to be 59.2 kJ g-1 and Student B determines the value to be 54.7 kJ g-1 then Student B’s results are said to be more accurate than student A’s results. If the accepted value for the heat of combustion was not known, comments on accuracy would not be valid.

The relationship between accuracy and precision

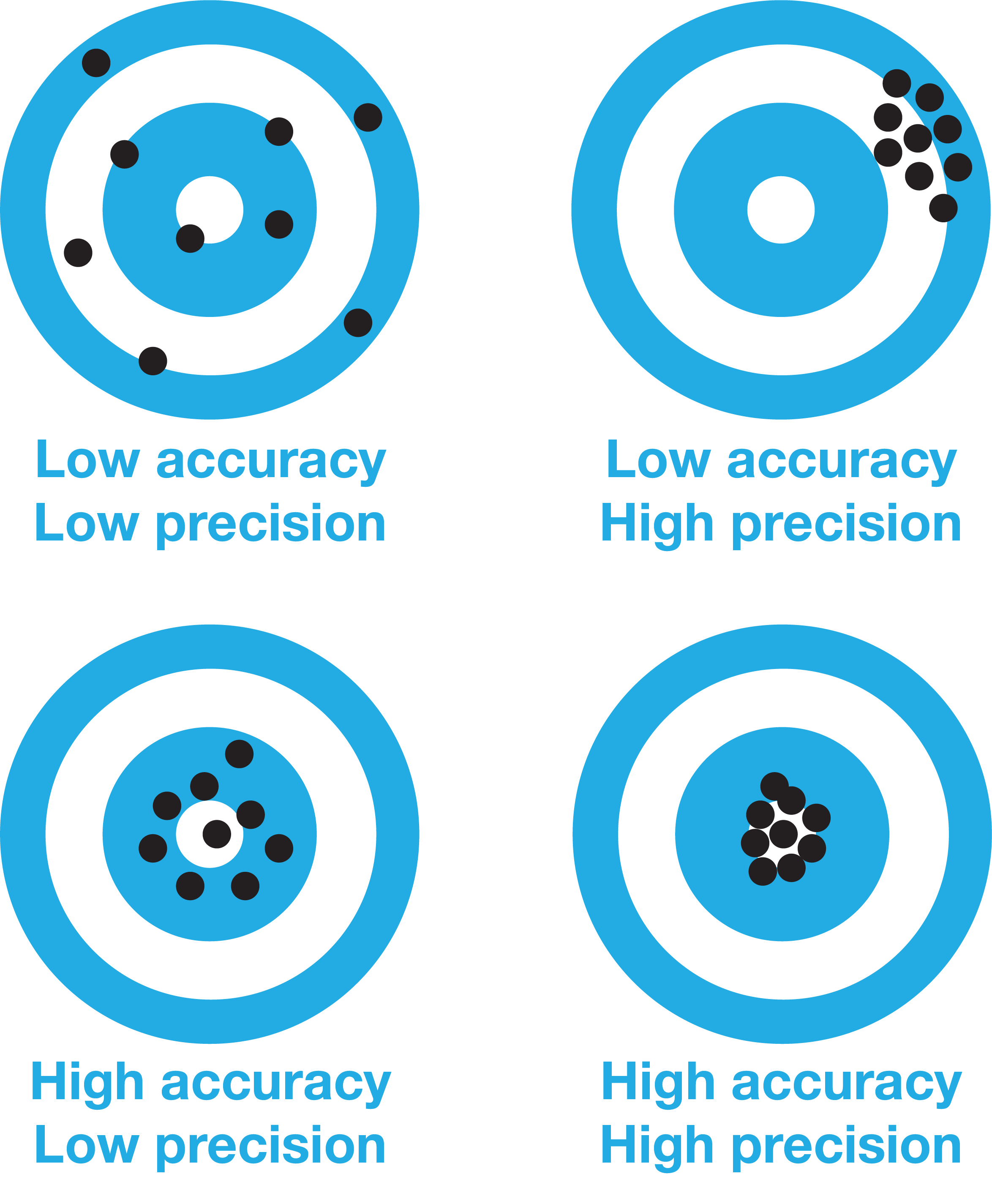
A set of measurement results that are close to the true, or accepted, value of a quantity are said to be accurate. A set of measurement results that are close to each other are said to be precise. For example, if a given substance was weighed five times, and a mass of 2.70 g was obtained each time, then the experimental data are precise. 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.

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 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 diagram below shows four representations of accuracy and precision.



The bullseye (centre of the concentric circles in each of the four representations) indicates the true value, while the nine black dots in each representation shows measurements.

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.

Repeatability refers to the precision, or 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 experiment several times in order to confirm their findings; this practice is also important for VCE Environmental Science students wherever practicable.

Reproducibility refers to the degree of agreement between the results of experiments 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 methodologies that other scientists can modify and enhance.

In the context of an experiment, repeatability measures the variation in measurements taken by a single instrument, person or team under the same conditions, while reproducibility measures whether the results of an entire study or experiment by one researcher or research group can be reproduced by other independent groups.

Reliability

The *VCE Environmental Science 2022–2026 Study Design* does not refer to the ‘reliability’ of 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 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 measuring a height when the depth should have been measured, or misreading a scale on a thermometer as 35 oC rather than 25 oC, or forgetting to divide the diameter by 2 before calculating the area of a circle using the formula A = π r2. 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. For example, in reading an instrument where the needle fluctuates unpredictably, it is equally likely that the readings could be above or below a reading obtained if the needle was not moving. 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 electronic balance that has not been calibrated may result in 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 of a consistent magnitude, depending on the cause of the error. 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 are expected to distinguish between random and systematic errors, and to explain the effects on the quality of the data of these errors in their own and others’ experiments. For example, if measurements are taken with a plastic measuring tape that has been stretched over the years, 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 didn't match their expectations, they should try and speculate why, suggesting further experiments to confirm their ideas.

Uncertainty

All scientific efforts are directed towards reducing the degree of uncertainty in the world about observations, relationships and causes. All measurements are subject to uncertainty and may have many potential sources of variation, including human capacity in measurement, the nature of the instrumentation used and external factors such as air currents in the laboratory. This uncertainly extends to all inferences and conclusions that depend on uncertain measurements.

VCE Environmental Science requires only a qualitative treatment of uncertainty (no calculations). 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). 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.