THE LANGUAGE OF MEASUREMENT

Terminology used in school science investigations

Published by ASE Publications on behalf of ASE-Nuffield

The Association for Science Education

Nuffield Foundation

Introduction

The aim of this booklet is to enable teachers, publishers, awarding bodies and others to achieve a common understanding of important terms that arise from practical work in secondary science, consistent with the terminology used by professional scientists. This vocabulary underpins all empirical science and so is applicable not only to school science experiments but also to evaluating aspects of scientific claims made in the public domain.

The booklet does not cover a complete range of 'how science works' skills. For example, little of the terminology used when evaluating science-related media claims is included, e.g. for a Gateway 'Science in the News' task or a Case Study in Twenty First Century Science. Note also that presentation of data in this publication, e.g. in tables and graphs, is incidental and not of primary concern.

Structure of the booklet

The booklet contains a glossary of key terms, plus a selection of investigations illustrating the terms (shown in bold text and with a cross reference back to the glossary) being used in context. The examples, drawn from school level biology, chemistry and physics, are written in an idealised 'pupil voice', which shows possible end points for an able pupil. A commentary is offered below each section of pupil material in the example investigations and is distinguished by italic text. The commentary explains the terms being used and how these are appropriate. Some examples indicate ways that the investigation, or the terminology used, might be modified for use with younger and/or older pupils. It is useful to have the glossary to hand when you read the illustrative investigations.

The glossary

The glossary is presented in three Parts, each of which contains a group of concepts associated with one aspect of experimenting and measuring.

Part 1: Measurements
Part 2: Measuring instruments and measuring systems
Part 3: Designing and evaluating school science investigations

Within each part, the concepts are listed in order of meaning, rather than alphabetically. More basic terms are introduced first; later terms build from the earlier terms.

Meanings given to every term in the glossary have been carefully considered. (See 'The process leading to this publication' below.) Wherever possible, we simplify the language used by metrologists, yet avoid any oversimplification that could lead to confusion and misinterpretation. The booklet is aimed at adult readers with a good scientific understanding. It is not intended for a pupil audience.

Illustrative investigations

The purpose of the example investigations is to illustrate terms being used in the context of school level experiments. They are not intended as examples of teaching activities to be undertaken by pupils.

The 'pupil voice' used in the illustrative investigations indicates ideas and terms that can be discussed with pupils. It does not suggest what pupils should be achieving, and it should not be used as a guide for assessing pupil understanding of the principles of measurement and experimental design. Science teachers must judge for themselves what terms to introduce to their pupils and how to achieve progression through the secondary school years.

'Pupil voice' is shown as full column width, in normal text; the Commentary is indented slightly and shown as italic text.

Context and background

The scientific community has developed over time a specialised vocabulary for talking about practical exploration and investigation. Terms such as valid, accurate, precise, error, uncertainty – and many more – are frequently used in discussing the design and outcomes of practical enquiry. Like other examples of specialist terminology, its purpose is to enable people to communicate clearly and effectively with each other about their work and its outcomes.

School science

In school science, pupils are introduced to some of this vocabulary and the ideas it is designed to encapsulate and express. The booklet came to be written because there were problems with how this was done. There was no general agreement within the science education community on the terms that were used or their meaning. So we could find the same word being used to express different ideas, or two (or more) different words being used to express the same idea. It could also be unclear if a word like 'reliability' or 'accuracy' was being used as an everyday term, or as a technical term with a precise and specific meaning.

This became a more serious issue when pupils were being assessed on their understanding of these words, or their ability to use them correctly and appropriately in reports of their own practical work. Assessment of practical work, in some form, has long been a part of school science. Since the mid-1980s, this has involved assessing pupils' ability to plan and carry out practical investigations. The way this has been implemented has increased the prominence of a particular vocabulary for talking about aspects of investigation design, and the collection, analysis and interpretation of data. More recently, the emphasis on 'how science works' heightened the demand on pupils (and teachers) to be able to talk about these matters in a clear and consistent way.

One aspect of the situation that we faced in 2009 was the differences between awarding bodies in the terminology used to talk about investigative work. Some GCSE specifications included a glossary of terms associated with investigative work in science, but there were differences in the terms included and in how these were defined. And there were differences in whether (and how) pupils' use of these terms was examined in written papers or through coursework. With many specification-related publications there were corresponding differences in teaching materials from different publishers.

The result was confusion among teachers and pupils about how the special terminology associated with practical investigation was to be learned, when it should be used, what specific terms mean, and how they appear in coursework marking criteria. This was a particular concern for ASE and Nuffield, which publish educational materials that are intended to be useful for all teachers, whatever science course or specification they may be using.

The word 'reliability' has posed particular difficulties because it has an everyday usage and had been used in school science to describe raw data, data patterns and conclusions, as well as information sources. On the strong advice of the UK metrology institutes, we avoid using the everyday word 'reliability', because of its ambiguity. For data, the terms 'repeatable' and 'reproducible' are clear and therefore better. For conclusions from an experiment, evaluative statements can mention 'confidence' in the quality of the evidence.

Scientific enquiry

The core aim of science is to develop explanations for natural phenomena and events that can be seen to be grounded in empirical evidence. Hence scientific enquiry has a central concern about the quality of evidence and of explanations that are based upon it. This publication focuses on ideas about scientific enquiry that pupils should consider in their own practical work: Can I rely on the data when drawing a conclusion? Are uncertainties in the measurements small enough? Does the difference between one measurement and another reflect a real change in the thing being measured?

Error and uncertainty approaches to measurement

Historically, the objective of measurement was to determine an estimate of the true value of some quantity that was as close as possible to that single true value. The Error Approach to measurement, as it is called, is still commonly used in schools and colleges. Unfortunately, school pupils too often think 'error' implies a mistake they have made, rather than inherent variations which affect measurements.

In recent years, metrology institutes internationally have come to prefer the Uncertainty Approach. Here the objective is not to determine a true value as closely as possible. Rather it assumes that the information from a measurement only permits a statement of the dispersion (interval) of reasonable values of the quantity being measured, together with a statement of the confidence that the (true) value lies within the stated interval.

Both approaches assume that no mistakes have been made in performing the measurement. In both approaches, analysis of variation in data enables a responsive scientist (or pupil) to improve the design of the experiment and thus improve the quality of measurements or any result calculated from measurements.

The process leading to this publication

The issues outlined above came increasingly to the attention of the ASE 11–19 Committee and were discussed by them in autumn 2008. ASE contacted Nuffield in January 2009 to discuss the situation further. It was agreed to invite representatives of awarding bodies, QCA, Ofqual and the National Strategy, and the UK metrology institutes (as custodians of scientific 'best practice') to a meeting in April 2009. As a strategy for improving the situation, it was agreed that work should start from internationally agreed terminology and adapt it (i.e. simplify it, without deviation if possible) to school science needs. The main reference source therefore has been *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms*, VIM, 3rd edition, JCGM 200:2008. It was also agreed that the use of terminology appropriate to investigative practical work should not be examined out of context, and that any agreement on 'approved terminology' should apply throughout the United Kingdom.

Following this, a working group (for membership, see 'Acknowledgements' page 2) met on three occasions to prepare suitable documentation. A second meeting with awarding bodies and metrology institutes took place in September 2009, to discuss a draft of the booklet.

Final editorial decisions were then taken by ASE and Nuffield.

Glossary Language of scientific measurements and investigations

This glossary contains basic vocabulary used when discussing experimental design, measuring instruments, the validity and quality of data, or associated conclusions. The tables are in three parts:

1 Measurements. **2** Measuring instruments and measuring systems. **3** Designing and evaluating school science investigations.

Within each part, the concepts are listed in order of meaning, rather than alphabetically. More basic terms are introduced first; later terms build from the earlier terms. References within the table take you to examples showing each term being used in context.

Part 1: Measurements

Suggested school usage	Meaning ● notes	Example(s) showing this term used in context	VIM term
quantity	any property that can be given a magnitude by measuring, weighing, counting etc. For example, length, heart rate, electric current, volume of liquid, chemical concentration. (This is in contrast to a 'quality'.)		quantity
value (of a quantity)	number and reference together expressing the magnitude of a **quantity**. This can be a number plus unit (e.g. 3.0 A), number plus reference to a procedure (e.g. adhesion value of sticky tape), or number plus reference material (e.g. Mohs scale of hardness).	Flow rate in a stream (15) Using microtitration (30)	quantity value
quantity you are trying to measure	**quantity** intended to be measured. For example, length of a piece of string, temperature of a body, time for ten swings of a pendulum	Flow rate in a stream (15) Headstones (22) Brown paper towel (26) Using microtitration (31)	measurand
validity of a measurement	a measurement is 'valid' if it measures what it is supposed to be measuring ● it depends on both procedure and instrument	Potassium salts (25) Brown paper towel (26)	
measurement result	**value** attributed to the thing being measured (measurand), reported at the end of a measurement process. Ideally expressed as a measured quantity value and a measurement **uncertainty.**	Speed of sound (21) Brown paper towel (30) Using microtitration (32)	measurement result
measurement, measured value	**value** representing a **measurement result**, which can either be obtained from a single act of measurement or by averaging individual measured values	Flow rate in a stream (16) Speed of sound (20,21) Brown paper towel (28) Using microtitration (30,31) Car tyres (35)	measured quantity value (measured value of a quantity)
true value	**value** that would be obtained in an ideal measurement ● considered unknowable ● in the special case of a fundamental constant, the constant is considered to have a true value	Speed of sound (20) Headstones (23) Using microtitration (31,32)	true quantity value (true value of a quantity, true value)

Suggested school usage	Meaning ● notes	Example(s) showing this term used in context	VIM term
accuracy	a **measurement result** is considered accurate if it is judged to be close to the **true value**. A quality denoting the closeness of agreement between a measured value and the true value of a measurand. ● not quantifiable (more or less accurate) ● property of a single result, which is influenced by both random and systematic errors	Speed of sound (19,21) Headstones (23) Brown paper towel (27,28) Using microtitration (31,32)	measurement accuracy, accuracy
precision	a quality denoting the closeness of agreement between (consistency, low variability of) measured **values** obtained by repeated measurements ● depends only on the extent of random effects – it gives no indication of how close results are to the **true value** ● a measurement is 'precise' if values cluster closely ● can be expressed numerically by measures of imprecision (e.g. standard deviation)	Using microtitration (32) Car tyres (35)	measurement precision, precision
repeatability (when comparing results from the same pupil or group, using same method and equipment)	**precision** obtained when **measurement results** are produced in one laboratory, by a single operator, using the same equipment under same conditions, over a short timescale. A measurement is 'repeatable' in quality when repetition under the same conditions gives the same or similar results e.g. when comparing results from the same pupil or group, using the same method and equipment.	Speed of sound (20,21) Headstones (23) Potassium salts (25) Brown paper towel (27,29) Using microtitration (32) Car tyres (35)	measurement repeatability
reproducibility (when comparing results from different pupil groups, methods or equipment – a harder test of the quality of data)	**precision** obtained when **measurement results** are produced by different laboratories (and therefore by different operators using different pieces of equipment) A measurement is 'reproducible' in quality when reproducing it under equivalent (but not identical) conditions gives the same or similar results e.g. when comparing results from different pupil groups, methods or equipment – a harder test of the quality of data.	Flow rate in a stream (19) Speed of sound (21) Headstones (23) Brown paper towel (29) Using microtitration (32) Car tyres (35)	measurement reproducibility
uncertainty	interval within which the **true value** can be expected to lie, with a given level of confidence or probability, e.g. 'the temperature is 20 °C ± 2 °C, at a level of confidence of 95%'. ● whenever a measurement is made, there will always be some uncertainty or doubt about the result obtained ● can be expressed in terms of standard deviations or other estimate of spread (e.g. range of values obtained, or interquartile range). ● sources of variation in the data collected include contributions from both random and systematic effects ● Uncertainty can also be estimated by understanding the instruments used (e.g. typically the uncertainty might be estimated as ± half the smallest scale division) and what effect any outside perturbations might have (e.g. a lab bench jiggling).	Flow rate in a stream (16) Speed of sound (20,21) Headstones (22,23) Brown paper towel (26,27) Using microtitration (31,32,33) Car tyres (33,34)	measurement uncertainty

Suggested school usage	Meaning ● notes	Example(s) showing this term used in context	VIM term
measurement error	the difference between a **measured value** and the **true value** of a physical quantity being measured, where a true value is thought to exist. ● It is important not to confuse the term 'error' with a 'mistake' in measurement, or with **uncertainty**. ● Whenever possible scientists try to correct for any known errors: e.g. by applying corrections from **calibration** certificates. ● Any error whose value is unknown is a source of **uncertainty.**	Speed of sound (19)	measurement error
random error	component of **measurement error** due to measurement results varying in an unpredictable way from one measurement to the next. ● Random variation is present when any measurement is made, and cannot be corrected for. The effect of random variation can be reduced, however, by making more measurements and reporting the mean. ● Random variation arises from uncorrelated effects of factors which are not controlled, e.g. electrical noise. ● Random errors cause an element of **uncertainty** in the result known as 'random uncertainty'.	Using microtitration (33)	random measurement error
systematic error	component of **measurement error** due to measurement results differing from the **true value** by a consistent amount each time a measurement is made. ● In some cases a systematic error leads to a constant offset (a fixed amount in one direction). In other cases, systematic effects are not constant but follow a pattern, e.g. dependence on prevailing temperature. ● The magnitude (and direction) of systematic effects determine the measurement bias in values obtained. Systematic effects can be caused by influence of the environment, methods of observation or instruments used. ● It may be possible to reduce or remove systematic errors if their causes can be understood and corrected or removed. For example, checking the zero reading of an instrument during an experiment as well as at the start, to ensure that it has no **zero error.** ● Even though it may be constant, a systematic error can remain unknown. There are many historical examples of measurement results that were later shown to contain systematic error. ● Unknown systematic errors cause an element of **uncertainty** in the result known as 'systematic uncertainty'.	Flow rate in a stream (16) Speed of sound (19,21) Headstones (22) Brown paper towel (27) Using microtitration (34) Car tyres (35)	systematic measurement error

Illustrative investigations

These investigations illustrate glossary terms being used in context. Glossary terms are shown in bold text, with a cross reference back to the appropriate glossary page to help you refer back to the glossary as you read each investigation.

The examples are written in an idealised 'pupil voice', which shows possible end points for an able pupil. A commentary is inserted at intervals within the pupil writing and can be identified by the use of italic text. This explains the terms being used and how these are appropriate. In some examples, it also indicates ways that the investigation, or the terminology used, might be modified for use with younger and/or older pupils.

Does the flow rate in a stream affect the numbers of freshwater shrimp?

About this example

In any investigation, it is important to use the underpinning procedural ideas to evaluate the quality of the investigation and the resultant data.

In fieldwork investigations many of the procedural ideas are the same as those used in lab-based investigations. Fieldwork introduces pupils to additional procedural ideas, many of which are important for understanding the evidence in many real-life situations.

In many biological investigations, an important issue to consider is large inherent variation. Collating class data is often a practical way to address this. The size of the sample is something that needs to be emphasised.

Large data sets, sometimes with measurements of several variables, are often needed. This increases the procedural complexity for pupils. Using a few values of a categorical variable (say, 'upstream' and 'downstream' of a source of pollution) instead of many values of a continuous variable (such as readings along the length of a stream) helps to reduce the amount of data and hence the procedural complexity.

Introduction

The flow rate in a stream might affect the particle size of the substrate (e.g. silty, small stones, bigger stones etc.); the availability of food since organic matter will be in suspension; and also oxygen concentration. Therefore flow rate, which varies in different parts of the stream, may well affect the distribution of aquatic species.

We will carry out a survey in the Mill Race Stream to see if the numbers of freshwater shrimp, my **dependent variable** (12), are affected by the flow rate (one of many other **variables** (12)). We have decided to select just three sites with different flow rates in shallow water.

The pupils recognise that flow rate is associated with these biologically important variables.

By naming categories, 'Substrate' is being considered as a categorical variable. It could be measured by particle size (a continuous variable).

This survey attempts to isolate the possible effect of one *variable on another. This is easier for pupils to cope with than investigations with more independent and/or dependent variables.*

Flow rate can also be measured along a transect, with kick sampling occurring at the same points as the flow rate is measured. This would involve far more data collection *and analysis, which would increase the procedural complexity.*

Although causality cannot be demonstrated in contexts where other variables may be affecting the relationship, the use of the term **dependent variable** (12) *can be justified in this context since it is clear that the numbers of shrimp are affected by other variables rather than vice versa. In other correlational studies where it is unclear which variable is affected by which other, then it is more appropriate not to identify variables as being independent or dependent.*

In ecological surveys we cannot change the values of any **control variables** (12) so cannot fix them at a constant value. We will have to make our investigation as valid as possible by monitoring the effect of other variables and checking that they are as similar as possible.

In a lab, where the values of variables can be manipulated by the pupil, the **validity of the experimental design** (12) *is ensured by keeping the values of control variables constant. In many biological contexts such manipulation is not possible so validity is increased by ensuring that comparisons of variables are made where the effects of other key variables are similar.*

The other key variables that we know may affect the distribution are pH, temperature and light. We will select sites where these are likely to be similar and will monitor these variables. Since all the sites are in the same stream, we can assume that the type of vegetation surrounding the stream, or being swept down from upstream, will affect all the sites in the same way.

The pupils control for the effects of these other variables that might affect the number of shrimp (the dependent variable) by ensuring that all sites are affected in a similar way.

Type of vegetation can be described in words: it is a categorical variable.

We can only do this data collection on one day (20 March) and at one time so we won't know how the time of day or year affects any of this.

The pupils are aware that the investigation is valid only under the specified conditions.

Instruments

All the variables need to be described, counted or measured.

Categorical variables have **values** (8) *that are labels, e.g. names of plants, descriptions of substrate or, as in this case, are labelled with numbers.*

Continuous variables can have values that can be given a magnitude either by counting (as in the case of the number of shrimp) or by measurement (e.g. light intensity, flow rate etc.).

We'll measure the flow rate using Poohsticks, timing how long the stick takes to travel 1 m at each site. It's easiest to show this as the time for it to travel one metre rather than as an actual rate. We know that there will be lots of variation in each reading, due to eddies, problems in timing etc., so we will repeat the Poohstick test six times at each site to get an average flow rate. We are only allowed old stop clocks. Some have a **zero error** (11); the needle won't go back to the start each time, so we will do a **zero adjustment** (11) by noting the starting **indication** (11) and subtracting this from the indication when we stop the clock.

Since flow rate at each site varies, the pupils attempt to capture the variation by taking a **sample** (12) *of repeated readings of the variable. The variation is due to random effects.*

The pupils recognise that they are not actually calculating the flow rate. The **quantity they are trying to measure** (8) *(the measurand) is clearly defined.*

The pupils attempt to correct the values on the clock because it does not start at zero.

The number of shrimp will be counted in a kick sample. One person will stand at each site, upstream from a net and will shuffle their feet for 1 minute to dislodge any animals into the net. This will be repeated twice at different places at each site to get an average reading. The sites aren't big enough to take any more repeated readings. The net will be emptied into a tray and all the freshwater shrimp counted before returning them to the stream. We can be confident that our identification and counting is correct – they're easy to see – but there will be measurement **uncertainty** (9) due to the kick sample.

Kick sampling dislodges animals in the substrate. The pupils are aware that variations in technique may affect the resultant data. Repeated readings attempt to estimate this variation. But since a kick sample cannot be made on the same spot, and the area is quite small, there are limitations to the number that can be taken. Pupils need to be encouraged to consider this when planning and also when they make their claim; their sample cannot be very large but may still be representative of the population of shrimps in that area.

Kick sampling may cause a **systematic error** (10) *since on the data since some organisms may be less easily dislodged than others. This may bias the results.*

In many biological investigations there is a lot of inherent variation. Therefore getting sufficient repeated readings in the sample to represent the variation is important. Large sets of readings are often needed, which is why, in biology, classes often work together and collate their results.

Water pH, temperature and light can be recorded with an electronic sensor at each site. Temperature and light intensity vary during the day and during the year, so this will be just a snapshot. The probe is **calibrated** (11) by the technician before we use it. We leave the probe in the water for three minutes before taking each reading.

Many pupils think that an instrument with a digital readout provides good quality data. Pupils should be encouraged to consider the quality of any instrument they use.

Electronic probes (and many other instruments, such as mercury thermometers) must be left to 'settle' before a reading is taken. This is due to their **response time** (11).

Procedure

We selected three different safe sites to work in the Mill Race Stream.

We worked in groups. Once we'd checked that pH, temperature and light were about the same at each site, then we measured the flow and did two kick samples at each site.

We recorded our data in a table.

Tables can be used as organisers for data collection. Pupils should be encouraged to design a table using the variables they have identified for the investigation.

The data

Site	Temperature in ^{o}C	pH	Light (W m^{-2})
1	8.00	6.70	584.00
2	8.00	6.86	605.20
3	8.00	6.86	600.93

The sites are numbered in this example. These numbers are just labels; in effect, values of the categorical variable 'sites'. Pupils need to be aware that numbers used as labels cannot be added, divided etc.

The **measured values** (8) of these key variables are similar at each site.

Similar values for key variables increase the validity of the relationship between flow

Site	Time in seconds to travel 1 m	Number of shrimp in 1 min kick sample
1	52	15
1	45	11
1	71	
1	42	
1	49	
1	80	
2	2.53	7
2	1.43	3
2	1.34	
2	2.02	
2	1.32	
2	1.55	
3	No movement	37
3	No movement	23
3	No movement	
3	No movement	
3	No movement	
3	134	

and shrimp numbers; i.e. making this a **valid conclusion** (12).

Site 3 is a pool. It has very little flow. When we dropped the Poohstick it stayed still. The final reading may have been affected by splashing further up stream, so we ignored it.

A mean has been calculated, after repeated readings for both the variables have been taken, since the values varied.

Site	Mean flow (time in seconds to travel 1 m)	Variation in flow: range	Mean number of shrimp	Variation in shrimp numbers: range
1	56.5	42–80	18	11–15
2	1.7	1.32–2.53	5	3–7
3	No flow	No flow, so no variation	30	23–37

The **range** (12) *is shown to give an indication of the variation, but the limited number of kick samples means that the data must not be over interpreted. It is not consistent but we don't know whether this is due to variations in the number of shrimp at each kick sample or the variations in the technique.*

Pupils should be encouraged to consider whether the uncertainties are small enough, to be convinced that any perceived differences are real.

An explanation was found for the anomalous data point and it can sensibly be ignored.

To show the patterns in the data, we rearranged the data.

Site	Mean flow (time in seconds to travel 1 m)	Variation in flow: range	Mean number of shrimp	Variation in shrimp numbers: range
2	1.7	1.32–2.53	5	3–7
1	56.5	42–80	18	11–15
3	No flow	No flow, so no variation	30	2337

Patterns can be more clearly seen if the data is ordered in tables.

We have data from a range of values of flow: from a still pool with no flow to a rate where the Poohstick moved 1 m in about 2 seconds. Although flow is a continuous variable, we only have three values with large **intervals** (12) between some of them. We could have shown these on a line graph but we'd have to be very cautious since we have little confidence in the data and the relationship is still uncertain.

However, it is difficult to plot a line graph with 'no flow', so we have decided to base our claims just on the tabulated data.

The continuous independent variable, flow rate, has been selected with a large range but the relationship is hard to determine both because of the quality of the data and the relatively few intervals selected.

A bar chart is used to show data from an investigation with a categorical independent variable and a continuous dependent variable.

A line graph enables interpolation and extrapolation to see the relationship between two continuous variables. However, the pupils identify that they have very limited data to give confidence in such a relationship.

Values on the axes of a line graph are integers. 'No flow' could not be plotted.

Evaluation and conclusion

We cannot claim that flow causes the variation in shrimp numbers. Not all conditions have been isolated, so we can only show that there is an association.

The difference between a causal relationship and an association or **correlation** (13) *is important.*

The measurements of the control variables suggested that conditions were similar on this day and time at each site. However, we do not know whether they would be similar in, say, the middle of winter. Since shrimp can be long lived, it may be one of these factors at a different time of year or other factors that are affecting their numbers.

All of the measurements show variation. Having only two kick samples at each site is a problem: we don't know much about the variation. But the numbers at each site are quite different so perhaps the differences are real.

In the Mill Race Stream the evidence suggests that more freshwater shrimp are found in slower moving water than fast, at least on the day we took our samples.

The claim is qualified and delimited.

If we had set up an 'artificial stream' in a lab, we may have been able to determine if flow rate caused there to be different numbers of shrimp. But identifying and isolating all the many variables would have been difficult so it would be hard to ensure such an 'artificial stream' was valid.

The pupils make a qualitative judgement of the **degree of confidence** (12) *they have in the claim.*

Since the conditions will be different in other streams and under other conditions the details are unlikely to be **reproducible** (9) *but the trend may be.*

Ecologists are sometimes able to isolate variables better under lab conditions, and so establish causation.

Measurement of the speed of sound

About this example

This is a standard experiment to derive a value for the speed of sound. Because there is an accepted value for the speed of sound at different temperatures, it is possible to have meaningful discussion about accuracy. The apparatus used claims microsecond resolution, so the pupil feels that it might be possible to get a very accurate determination of the speed of sound.

Two microphones are used: one starts the timer when the rising input reaches a certain value and the other stops the timer when the rising input reaches a certain value. To make a sharp pulse of sound a hammer strikes a small metal plate.

Repeating measurements enables comments about consistency and uncertainty to be made. Repeating the experiment enables comments about confidence in the derived value for the speed of sound to be made.

Introduction

I intend to measure the speed of sound using the microsecond timer. I thought that if I could measure time to a microsecond I could get a very **accurate** (9) value for the speed of sound.

Here the term accurate is being used to signify how close to the **true value** (8) *a measurement can be made using an instrument that appears to provide considerable* **resolution** (11).

Procedure

The time for sound to travel will be **proportional** (13) to the distance between microphones. I set up the microphones a metre apart and decided to make ten measurements of the time taken for a sound pressure wave to go from one microphone to the other. I measured the distance to one millimetre.

The pressure wave hits the first microphone and starts the timer, which stops when the pressure wave reaches the second microphone. In this way any **systematic error** (10) in the electronics is eliminated.

The pressure wave is made by striking a metal plate with a hammer in the same place along the line joining the microphones.

The pupil is concerned to minimise any **measurement error** (10). *Delays in the circuitry to start and stop the timer are assumed to be equal so that any systematic error caused by, for example,* **response times** (11) *of the microphones, are eliminated. It is by no means certain, however, that the microphones are identical and so ideally the experiment should be repeated with the microphones exchanged.*

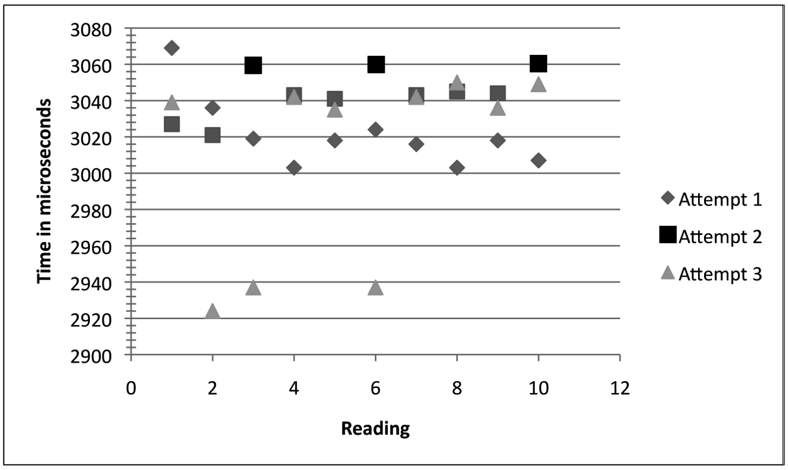
Results

Attempt 1	Attempt 2	Attempt 3
Time in microseconds		
3069	3027	3039
3036	3021	2924
3019	3059	2937
3003	3043	3042
3018	3041	3035
3024	3060	2937
3016	3043	3042
3003	3045	3050
3018	3044	3036
3007	3060	3049
Mean		
3016	3044.3	3041.8
Speed m/s		
331.56	328.48	328.75

All four figures from the timer have been quoted and the speed calculated to five figures.

I repeated the experiment twice more to see how **repeatable** (9) my **measurement results** (8) were; three attempts in all. I could see that in each attempt the time varied quite a lot.

After plotting the data on a graph, I judged one reading in Attempt 1 (3069 μs) and three readings in Attempt 3 (2924 μs, 2937 μs, 2937 μs) to be **outliers** (13). These are shown with grey shading in my table.



I can see that the **uncertainty** (9) in the time is definitely more than 10 microseconds. The **range** (12) in each experiment is 29, 39, and 15 microseconds. The uncertainty in each attempt might be as much as ± 30 microseconds but the uncertainty in the mean might be less (half), at about ± 15 microseconds.

The result is calculated from a set of **measurements** (8).

The pupil is reconciling the microsecond **indications** (11) *on the measuring instrument with ideas about uncertainty. If pupils have an understanding of standard deviation then it could be used here. Repeated readings show how close the agreement is between individual measurement values in one measurement set and how repeatable the result* *(speed of sound) is by repeating the experiment (the ten measurement values).*

The pupil is showing that he/she knows that taking the mean of a reasonably large number of data will reduce the uncertainty for the mean.

It would be good to find out how **reproducible** (9) *the result is by comparing results with other classes using the same apparatus.*

The accepted value of the speed of sound at 12 °C is 339 m/s. At this speed, sound only takes about 3 microseconds to travel 1 mm so I know that the distance measurement has been made accurately enough, since the variation in times is much bigger than 3 microseconds.

The pupil is implying that even if the true value is slightly different to the measured value it is unlikely to make a big contribution to the uncertainty in the result if it changed by ± 1 mm.

I have quoted my result to five figures, but because the **uncertainty** (9) in distance is about 1 mm, which represents an uncertainty in the speed of about ± 3 m/s, my three results become

(331 ± 3) m/s, (328 ± 3) m/s, (229 ± 3) m/s

So although the timer measures to a microsecond and my distance measurement is accurate it does not make sense to give more than three **significant figures** (13) in my result because of the uncertainties in the time measurements.

I don't know why my time measurements varied so much. Maybe the microphones respond differently depending on the sharpness of the strike so that the timer is triggered to start and stop at slightly different times. However, my experiment is repeatable as all three attempts are within each other's range of uncertainty.

The pupil has shown that the data are sufficiently consistent to provide a repeatable result.

The pupil is attempting to explain why the individual **measured values** (8) *vary as well as making a comment on the repeatability.*

The microphones themselves may be different and so really I should have reversed them or changed ends for the strike, repeated it and then taken the mean of each direction.

I have shown that small changes in the distance between the microphones will have little effect on the times and the temperature in the room did not change enough to make a difference either.

I am **confident** (13) that my **measurement result** (8) is repeatable but generally lower than the accepted result so there must be a **systematic error** (10) caused either by the method or the apparatus.

The signal drops in amplitude as it goes from the first to the second speaker; this may account for the timer stopping slightly later, resulting in a lower than expected value for the speed of sound.

The pupil is able to determine that there is a systematic error *caused by the method or the apparatus. The difference between the results and the true value is greater than can be accounted for by the measurement uncertainty, which suggests an additional systematic effect. Attempting to explain the bias can point to different methods or the use of different or modified apparatus.*