

Accuracy and reliability

Background

A pupil working at:

Level 4

- knows that the accuracy of measurement is limited by the equipment used (e.g. mm on a ruler)
- repeats results but doesn't know why
- confuses accuracy with reliability (has no real understanding of what 'reliable' means)

Level 5

- begins to appreciate that the range of data collected can affect the reliability of evidence
- understands that repeating results checks reliability
- knows when and when not to repeat readings

To move pupils from level 4 to level 5

1. Look at a range of measuring instruments and make judgements on how accurately you can measure with them.
2. Use model data to encourage judgements on accuracy.
3. Use model data (tables and/or graphs) to explicitly show that a limited range of data can sometimes yield **unreliable** evidence.
4. Use model data where results are repeated. Make sure that in the model one reading doesn't fit (is anomalous). Explicitly point out that when readings are the same or similar they support one another and are therefore considered **reliable**.
5. Consolidate pupils' understanding by letting them carry out a simple experiment to demonstrate whether they are capable of collecting accurate and reliable results.

Suggested teaching sequence

The following activities could all be done in one lesson or split into separate activities and spread over more than one lesson.

1. To teach pupils to look at a range of measuring instruments and make judgements on how accurately you can measure with them

- Look at a range of measuring instruments, e.g. *range of measuring cylinders, range of Newton meters, rulers that measure to the nearest cm and mm, trundle wheel*, etc. Discuss what they are used to measure, and look at the scale to see what is the maximum the instrument will measure and what steps the scale goes up in.

- Pupils work in groups to discuss how accurately they could measure with each measuring instrument.
- Use sheet 5, 'Which measuring instrument?', and ask pupils to decide the most appropriate measuring instrument for each task. Other examples can be added, for example, to use a voltmeter, ammeter, CRO (for sound).

2. To teach pupils to make judgements about accuracy

- Use sheets 6 and 7, 'How accurate?', which could be converted to OHTs. The sheets contain data on the mass of magnesium before and after heating, and on pupils running 100 m, with questions for pupils to discuss. Use this activity to explicitly illustrate that sometimes evidence can be unreliable due to limitations of the selected measuring instrument and how accurately it is able to measure. There may also be differences between similar pieces of equipment, which you can demonstrate by asking pupils to measure the same item (e.g. temperature of boiling water, weight of a 100 g mass) on a series of measuring instruments.
- Discuss with pupils the meaning of 'accurate'. First, there is accuracy of the measuring instrument and pupils need to select equipment with the right degree of precision for an investigation – this links with the activity above. Second, it can be difficult to measure accurately due to the conditions, e.g. the height a ball bounces or the extension of a spring that wobbles.
- Use *AKSIS Investigations: developing understanding in scientific enquiry*, pp. 96–98.

3. To teach pupils that a limited range of data can sometimes yield unreliable evidence

- Look at sheet 8, 'Making cheese'. Ask pupils to discuss in pairs any simple conclusions that they can draw from groups 1, 2 and 3 – 'What do their results tell them?' This is a good opportunity for pupils to revisit fair testing and the key words **independent** and **dependent** variables by asking:
 - What is the **independent** variable?
 - What is the **dependent** variable?
 - What things (variables) should they keep the same?

Pupils also have the opportunity to practise describing the relationships between the variables. Make explicit that groups 1 and 2 would draw different conclusions to group 3 due to number and range of measurements. One aspect of reliability therefore is to collect sufficient evidence.

- Discuss with pupils the meaning of 'reliable evidence'. If results are reliable, then every time the experiment is repeated, the results are exactly the same. Sometimes reliable results can be inaccurate if the measuring equipment is faulty.

- Refer to *AKSIS Investigations: developing understanding in scientific enquiry*, pp. 99–101.

4. To teach pupils that repeat readings are useful to judge the reliability of results

- Use sheet 9, 'Stretching a spring' – again, this could be converted to an OHT. You may wish to have equipment available to demonstrate briefly what the pupils did in their experiment. Use this to explicitly illustrate when results support each other (are reliable) and when they are not. Discuss how close results need to be to be considered close enough to each other. This will be brought out when pupils discuss the table, where they have to decide whether results are reliable or not.
- Pose the question: 'What could we do about the unreliable result?' Suggestions should include repeating a third time, comparing this to the other two, identifying anomaly, judging which numbers to include in the average.
- Refer to *AKSIS Investigations: developing understanding in scientific enquiry*, pp. 65–68, on taking repeat readings.

5. To consolidate that pupils have understood 'accurate' and 'reliable'

Choose a simple experiment (friction is a good one) and ask pupils to collect results that they consider both accurate and reliable. You could give them a checklist or use discussion to generate a class checklist which will help them justify the accuracy and reliability of their results. For example:

- Have they selected the most appropriate measuring instrument?
- Have they measured as accurately as they can with the apparatus chosen?
- Have they repeated all readings twice, and repeated a third time only when a result seems odd?
- Have they collected sufficient evidence to be able to look for trends or patterns or to answer the question?

Relevant Key Stage 3 national test questions

- **2003 paper 1, tier 3–6 question 8**
Selecting equipment, predicting and considering results
- **2003 paper 1, tier 3–6 question 12; tier 5–7 question 2**
Considering results
- **2001 paper 1, tier 3–6 question 4**

Which measuring instrument?

What would you use to measure

1. your mass?
2. the diameter of your head?
3. the mass of a bean seed?
4. 10.5 g of rock salt?
5. the distance across the school field?
6. 1 cm³ of alcohol?
7. 250 cm³ of water?
8. the weight of an apple?
9. the change in temperature in the classroom over a weekend?
10. your body temperature?
11. the time it takes for your pulse to return to normal after exercise?
12. the time every 2 minutes – so that you can take the temperature of the water as it cools?
13. the amount of air in your lungs?
14. the pH of the pond water?
15. the amount of light under a tree?
16. the height of a plant every day?
17. the mass of a slice of cucumber before and after it is dried?
18. different volumes of water from 10 to 100 cm³?
19. the length of a woodlouse?
20. the brightness of a light bulb?
21. the time for a spinner or parachute (made in the classroom) to fall?
22. how far an object slides on different ramps?
23. the length of dandelion leaves?
24. 0.5 g of copper sulfate?
25. the number of seconds a pendulum swings for?

How accurate?

1 Burning magnesium

Two groups of pupils used balances to find the mass of magnesium powder at the start of an experiment and then after they had burnt it in air.

Group 1

Mass of magnesium at the start = 1.2 g at the end = 1.2 g

Group 2

Mass of magnesium at the start = 1.19 g at the end = 1.24 g

Some questions about these results

- Do you think the two groups used the same balance? Can you explain your answer?
- Do the results tell you the same thing? Do they show any change in the mass of magnesium?
- Why did group 1 get the same readings before and after heating the magnesium?
- Why did group 2 get different readings even though they were doing the same experiment?
- Which group's results gave the most accurate result and why?

2 Running 100 metres

Four pupils decided to use three different stop-watches for the same race. They wanted to see if the type of stop-watch they used to measure the time it took them to run 100 m really made a difference to their results. This is what they found out.

Results 1 – 100 m dash

<i>Pupil in class</i>	<i>Time (seconds)</i>
Sian	14
Chris	14
Des	15
Phil	14

Results 2 – 100 m dash

<i>Pupil in class</i>	<i>Time (seconds)</i>
Sian	14.4
Chris	14.1
Des	14.6
Phil	14.1

Results 3 – 100 m dash

<i>Pupil in class</i>	<i>Time (seconds)</i>
Sian	14.42
Chris	14.13
Des	14.62
Phil	14.08

Some questions about these results

- What do you think were the differences between the three stop-watches?
- Who ran the fastest in each race?
- Did the type of stop-watch make any difference to deciding the winner?

- Who came second in each race?
- Did the type of stop-watch make any difference to deciding who came second?
- How did you decide who came second?
- Which do you think was the best stop-watch to use for this investigation and why?

Making Cheese

Farmer Stilton makes cheese on his farm by reacting milk, collected from his cows, with a chemical called Maxiren. He wanted to know the answer to this question: 'How does the temperature of the milk affect the amount of cheese made?'



Three groups of Year 9 students did an experiment to find out.

Here are their results.

Group 1 took two readings

Temperature of milk (°C)	Mass of cheese (g)
20	1.4
40	2.9

What do their results tell them?

Group 2 took four readings

Temperature of milk (°C)	Mass of cheese (g)
10	0.5
20	1.4
30	2.2
40	2.9



What do their results tell them?

Try to sketch the shape of the graph of these results.

Group 3 took eight readings

Temperature of milk (°C)	Mass of cheese (g)
10	0.4
20	1.2
30	2.3
40	3.0
50	2.4
60	1.5
70	0.7
80	0.0



What do their results tell them?

Try to sketch the shape of the graph of these results.

Stretching a Spring

Pupils hung different masses on the end of a spring and measured how much the spring stretched.

Three groups did the same thing, using similar equipment. This is what they found.

Mass (in grams)	Group 1	Group 2	Group 3
0	2	2.4	2.4
10	3	3.1	3.0
20	4	4.1	3.6
30	4	3.9	4.1
40	5	4.5	4.5

- What was the general trend as more masses were added to the spring?
- Is that the same trend for every group?
- Look at the results for each mass across the three groups. If the results from all three groups are close then we could say that the results are reliable. Use the table below and decide whether the results from all three groups are reliable or unreliable for each mass and why you think that.

Mass used	Tick for reliable; cross for unreliable	Why did you make that decision?
0		
10		
20		
30		
40		

- How do you think you could make the results from this investigation more reliable? Can you explain your answer?
- Do any of the results seem odd? How do you think you could find out whether the result is reliable?