Pupils' investigations annotated to exemplify points made in Roberts & Johnson (2015) Understanding the quality of data: a concept map for 'the thinking behind the doing' in scientific practice, *The Curriculum Journal*, 26:3, pp. 345-369

Ros Roberts **Jan. 2016**

Based on Gott, R., Foulds, K., Roberts, R., Jones, M., & Johnson, P. (1999). *Science Investigations:3*. London: Collins Educational, (pp. 63-74).

Introduction

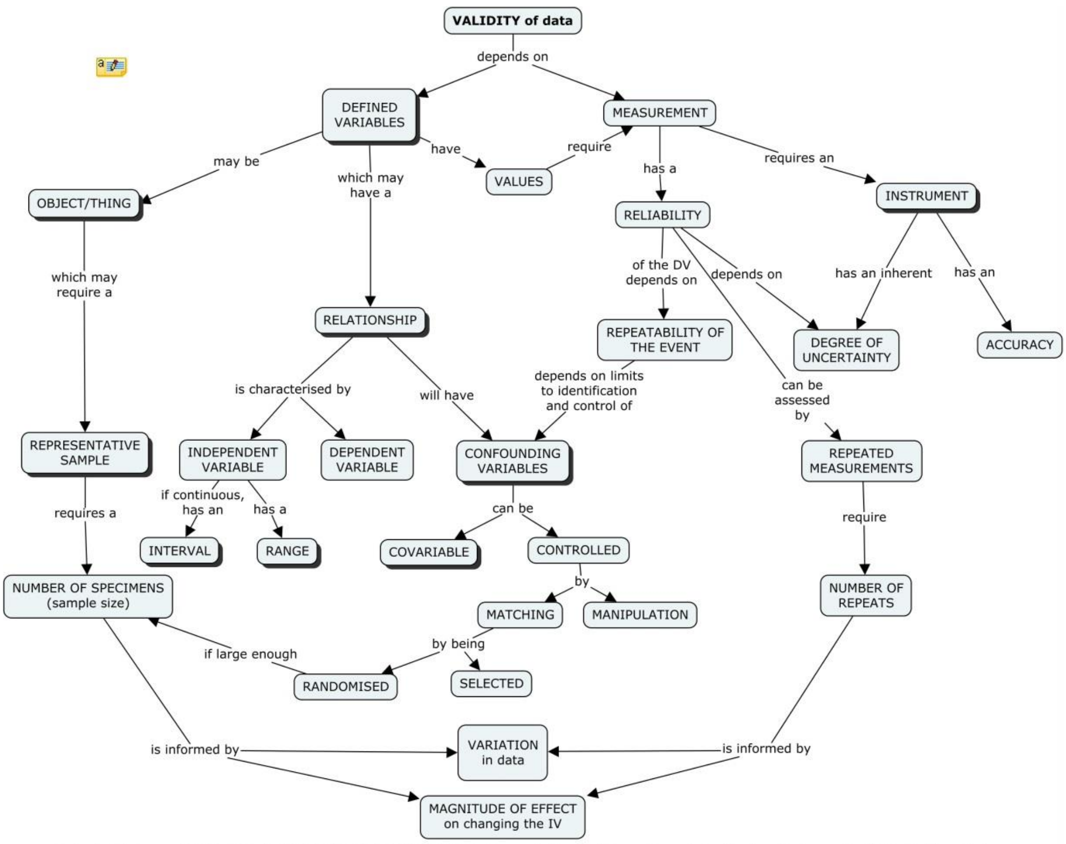
The accounts of the investigations, in conjunction with the questions at the end, are designed as activities to raise issues and address unexplained reasoning behind decisions (tacit knowledge / thinking). These pupils' accounts are not 'perfect examples' of executing investigations.

The investigations illustrate different approaches to collecting evidence and how the 'thinking behind the doing' of the concept map can be applied in such contexts.

> The conceptual overview represents a network of intricately linked ideas, and decisions when investigating are based on nuanced application of these ideas, involving mental juggling as juxtapositions and contingencies are considered according to context. In terms of validity, there is no distinction between approaches (such as an 'experimental approach' or an 'observational approach') to finding patterns in data (Cleland, 2002). No one approach is privileged over another; the key issue is what is appropriate depending on the circumstances, as illustrated … Of itself, the map embodies the realisation that 'there is no single set or sequence of steps followed in all investigations' (Lederman et al., 2014, p. 68). Roberts & Johnson (2015) p. 359.

The accounts illustrate the iterative nature of an investigation and the importance of trialling to make decisions about the quality of data.

The annotations exemplify points made in Roberts & Johnson (2015). (Ideas in **bold** in the annotations are concepts on the map).



Roberts & Johnson (2015). DOI: 10.1080/09585176.2015.1044459. Concepts directly informed by substantive knowledge are highlighted with a shadow.

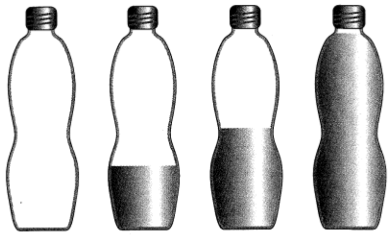
Complete investigations

Bottle tip

Some pupils were eating their picnic lunches on a sloping bank. They all had pop bottles with different amounts of drink in them. Some bottles wouldn't stay standing on the slope – they toppled over. Others stayed upright.

Does the angle of the slope at which the bottle topples over depend on how much pop is in the bottle?

You could use a large plastic pop bottle with a screw lid and water instead of pop.



You might need to think about these things before you start:

- what is a 'centre of gravity' and what does it mean?
- how will you create a slope?
- what will you need to try out in your trial run?
- what do you think might happen when you take all your results?

#1.Solving problems in science requires an understanding of both substantive ideas and ideas about evidence; they are inextricably linked (as shown on the concept map). This places great cognitive demand on the investigator. In this pupil investigation the substantive demand is relatively low so as not to detract from a focus on the ideas about the quality of data.

Poppy's investigation

How does the angle of the slope needed to tip a bottle over depend on how full the bottle is?

What are the variables?

independent variables — *amount of water*

dependent variables — *angle of slope*

type of bottle

surface

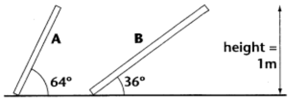
Trial run

I chose a plastic bottle and stood it on a short plank of wood. I lifted one end of the plank and saw when the bottle toppled over. I used the same bottle again, first full and then half full. It looked as if when it was half full I needed to lift the plank higher before it would topple, but I wasn't too sure about that.

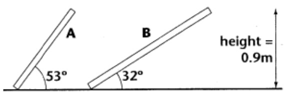
I could measure the height that the wood is lifted with a metre ruler.

I tried different lengths of wood and chose the longest I could find. Every time I lifted it up, the slope only changed a little. With a short one, lifting it made a big difference to the slope.

This diagram shows what I mean.



The planks are lifted to the same height. The short plank (A) has an angle of 64°, but the longer plank (B) only has an angle of 36°.



A small change in height changes the angle of the short plank (A) much more than that of the long plank (B).

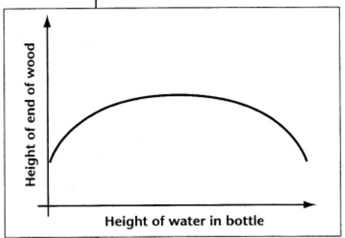
The surface of the wood made a difference. I didn't want the bottle to slide down. I tried several things but coarse sand paper pinned to the wood stopped it sliding.

What do I think will happen?

When it is empty I expect the bottle to tip over even when the slope is very gentle. But when I put water in it, it will make the centre of gravity lower. It's like when they make sure all the weight of a bus is low down so that it won't turn over. So I think it will stay upright until the slope is steeper (which is when the plank is lifted higher).

When it's full though the centre of gravity will go higher again. Then its like the bus being too full upstairs. I think that the slope which makes it topple will be less again.

So I think the steepness of the slope at which it topples might go down when it gets over half full. So the graph I draw might be something like this:



2

#2. The key **variables** affecting the **relationship** between **IV** and **DV** have been identified. In this simple context their identification does not rely on sophisticated substantive knowledge. All the variable values can be manipulated by the investigator.

#3. Trial runs characterise iterative working wherein decisions are made to establish the **quality of the data** (including the reliability of the DV, see #5). The investigator gets a feel for the relationship across the full **range of the IV**; and makes decisions to 'operationalise' the DV; and **manipulates** the values for the CVs. Decisions about one variable cannot be made without reference to others and determining their impact on the DV. The investigator avoids variation between types of bottle (shape and size) by selecting only one (and recognises this later, see #8). **Measurement** decisions are made to reduce **uncertainty** (but see also #8).

#4. Scientific theory enables predictions (hypotheses) to be made, which in turn may be tested by experimentation.

My plan

1 Measure the height of end of wood needed to tip the bottle when it's half full.
2 Repeat a few times to see how good the results are.
3 Decide how many times to repeat each measurement.
4 Measure height of wood when bottle topples for 6 different heights of water from empty (0 cm) to full (30 cm).
5 Find the average of each reading and plot a graph.
6 See what the results look like and see if I need do any more.

The results of my first 3 slopes

Height of water (cm)	Height of plank (cm)	Reading 2	Reading 3	Reading 4	Reading 5	Average
empty (0)	17	18	17	16	17	17
15	32	34	32	32	35	33
full (30)	18	19	18	18	17	18

#5. This establishes **a pattern over the range of the IV** in relation to the **scale of the variation** in the repeated readings (the **reliability** of the data). The investigator judges 5 repeats to be enough but does not explain why.

From these results I think I need to repeat each measurement 5 times and take an average. I now have to do the other slopes. Here's all my data.

Height of water (cm)	Height of plank (cm)	Reading 2	Reading 3	Reading 4	Reading 5	Average
empty (0)	17	18	17	16	17	17
5	42	42	44	43	44	43
10	41	32	40	41	39	
15	32	34	32	32	35	33
20	24	27	27	25	27	26
25	20	19	20	19	22	20
full (30)	18	19	18	18	17	18

#6. The reasoning that the variation in the IV in relation to changes across different values of the DV is behind the judgement to dismiss Reading 2 at 10cm as anomalous.

The value of reading 2, when there was 10 cm of water in the bottle, looked very different from the other readings. So I did it again and it tipped at 39 cm. I'll use this value to work out my average and ignore the odd one. The average height of the plank at which the bottle with 10 cm of water in tipped was 40 cm.

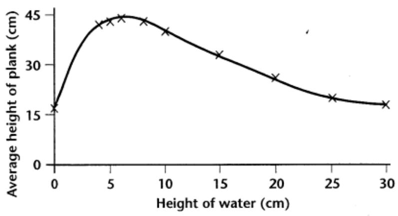
I decided to collect more data when there was 4, 6 and 8 cm of water in the bottle because it looks as if that might be where the curve turns down.

Height of water (cm)	Height of plank (cm)	Reading 2	Reading 3	Reading 4	Reading 5	Average
4	42	43	43	41	42	42
6	43	44	45	45	44	44
8	42	43	43	43	42	43

#7. Further data are collected at smaller **intervals of the IV** to help establish a pattern. This couldn't have been 'pre-planned'; it was in response to the quality of the data as it was collected.

In more traditional 'apparatus, methods, results' accounts, written up post hoc, the iterative working shown here would, by convention, have been presented as a more linear account.

What have I found out?



Do I believe my results?

On the whole I trust the results, at least enough to get a general relationship. The same bottle was used each time. I don't expect that I'd get exactly the same results with a different size or shape bottle. Each repeat gave a value very similar to the others, except for the one that I repeated.

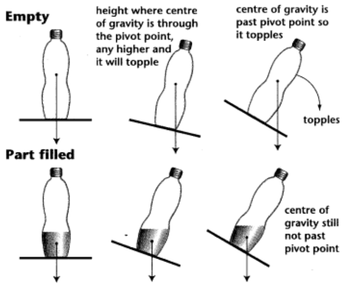
The extra readings with 4, 6 and 8 cm height of water in the bottle were useful. I think it shows me where the graph starts to turn down again but I should really have done some more repeats for the bottle with 4, 6 and 8 cm in it. The values are very similar and some of the values are even the same in different rows. I don't trust the reliability of this data as much.

I could also have done it with 7 cm of water as well.

I only measured the height of the water and the plank to the nearest centimetre. If I'd been more accurate my results would have been more reliable.

Explaining my results

The results fit reasonably well with what I expected to happen. The position of the centre of gravity is what's important. When the centre of gravity is low it is harder to topple because the bottle is not 'top heavy'. The bottle becomes top heavy when it is empty or full. When it is tipped, a top heavy bottle's centre of gravity is easily moved past the pivot point and it topples. If the centre of gravity of the bottle is low, when it's tipped it can go further before the centre of gravity goes past the pivot point. I've tried to show this in my drawing.



#8. The **validity of the data** depends on the **variation in the repeated readings** and the **magnitude of the effect of changing the IV**. Greater resolution of the **measurements** of both IV and DV would have increased the **reliability.**

The pupil attempts to reflect the quality of the data in the qualified claim.

Questions

1 Which were Polly's control variables? Highlight where she says they have been controlled.

2 The dependent variable is labelled as the angle of the slope. The measurement taken was the height the plank was lifted at one end. Can this be used as a measure of the angle of the slope? Explain your answer.

3 How else could Polly have measured the dependent variable?

4 Polly thought that five readings were enough so that an average could be worked out. What would Polly have had be take into account before she made this decision? Was it a good decision?

5 Explain in your own words why Polly ignored one of the readings and took another to use when calculating the average.

6 The values of the independent variable, the height of the water, were not evenly spread along the range. Explain why Polly chose these values.

7 After collecting her data, Polly collected another set of readings with 6 cm water in the bottle. Was this a good idea? Explain your answer.

8 When would the bottle have tipped if there had been 20 cm of water in it?

9 Underline in pen where Polly used ideas about friction in her planning.

10 Underline in pencil where you think the accuracy of Polly's experiment could have been improved.

#9. Pupils' understanding of the knowledge-base of evidence can be explored by targeted questioning, in just the same way as their substantive understanding can be assessed.

Complete investigations

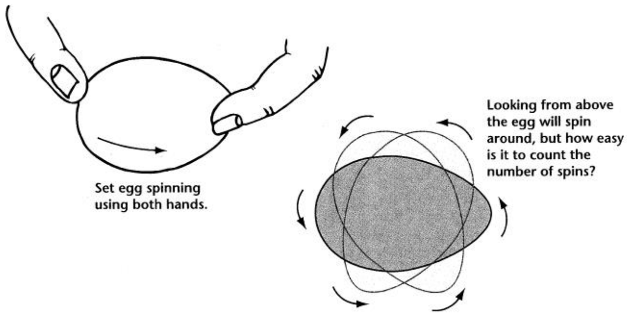
Eggs

If you spin eggs they will go on spinning for different times. Fresh eggs stop quickly and spin more slowly than hard-boiled eggs. If you'd got some different eggs, some cooked for different times and others raw, could you find out which was which without opening the shells?

Find out how the number of spins before an egg stops spinning depends on how long it has been boiled for.

#10. This investigation has very low substantive demand.

This investigation has data with more variation than in *Bottle Tip*. There is no variation in the sample tested (but it is noted between eggs); the cause of **variation is in the event** (the CV values) and variation in **measurement** of the DV.



You might need to think about these things before you start:

- do a trial run with a fresh egg – can you think of any reason why it might spin more slowly than a hard-boiled one?
- how will you set it spinning?
- what will you measure?
- what do you think might happen – what might your results look like?

1

Harbinder's investigation

How does the number of spins before an egg stops spinning depend on how long it has been boiled for?

What are the variables?

- *time egg boiled for*
- *size of egg*
- *shape of egg*
- *mass of egg*
- *number of spins before stopping*
- *time to stop spinning*
- *surface*

Trial run

Setting the eggs spinning wasn't easy. I tried using an old gramophone turntable but the egg rolled off. Even on a plate on the turntable it still rolled off. So I decided to spin the egg with my fingers on the kitchen table. I will need to try that out to see if I can do it the same each time.

I will count the number of spins until the egg stops, to the nearest quarter turn. I marked one end of the egg with a blob so I could see it when it span fast. I drew a line on the table and counted when the blob passed the line.

I will use the same egg all the time, and on the same surface. It might be different for different eggs so I will collect other people's data for to see if the pattern is the same.

What do I think will happen?

The number of spins will get more and more as the egg gets harder inside. I think this might be because when it is soft, the insides won't spin with the eggshell. It will flop around inside and slow it down.

My plan

1. *Pick a fresh egg.*
2. *Spin it with my fingers, by giving it a half-turn and count the number of spins until it stops.*
3. *Repeat a few times to see how repeatable the results are.*
4. *Put the egg in boiling water for a minute, take it out, hold it under running cold water for 2 minutes to stop it cooking and spin it again.*

SAFETY – I will need to wear gloves and use a spoon to hold the egg when I am taking it out of the water and cooling it.

5. *Then put it in the boiling water for another minute and do it again. Keep doing this for about 10 minutes until it should be hard-boiled. It shouldn't change after that.*

2

#11. Early decisions are made during trialling to help reduce the variation in the data. Only 1 egg is used thus avoiding variation in some of the identified CVs. Variable values can be **manipulated** to be kept more or less constant (with acknowledged variation in the values) to establish reasonable control of the **confounding variables**. Controlling the setting of the spin remains the biggest issue.

A table for my results – my trial run

Fresh egg – not boiled at all	
Trial run	**Number of spins**
1	10
2	9
3	7.75
4	6
5	15.25
6	9
7	17
8	15.1
9	15.5
10	16.75
Average	12.18

There is a lot of spread in the results from the fresh egg. I need to do some more trials at different boiling times to see if the spread is too big.

This is my data when the egg was hard-boiled.

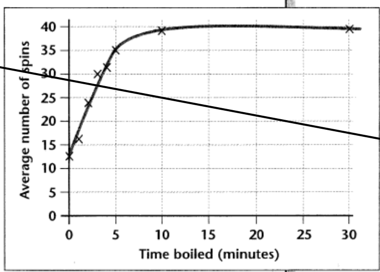
Hard-boiled (after 10 minutes boiling)	
Trial run	**Number of spins**
1	10
1	33.5
2	34.25
3	51
4	25.25
5	31.5
6	40.25
7	40
8	30
9	38.5
10	36
Average	36.03

There is quite a difference between the two sets of data. For my proper run I will get another egg and spin it 30 times for each amount of boiling, so that I can trust my average number of spins more.

These are my results:

Time egg boiled (minutes)	Average number of spins (from 30 repeats at each boiling)
0	13.1
1	16.5
2	23.7
3	30.4
4	31.5
5	35.1
10	38.8
30	39

This data will give me a line graph.



What have I found out?

The number of spins gets higher and higher as the egg gets more hard-boiled. At first it changes very quickly. After about 5 minutes it doesn't change as much. It must be getting hard inside by then.

Do I believe my results and why?

I don't think this method would be any good for testing to see how cooked an egg was. The results change a lot from one try to another. It isn't very reliable. I could use it to distinguish between the same egg when it was hard and raw but the data is not reliable enough for in-between.

I would need a much better way of getting the eggs spinning and I found it difficult to count it when it span very fast. If I'd recorded the egg spinning on video I could have slowed the tape down and counted more accurately.

I'd need to see whether cooling it in water after cooking really did stop the egg getting any harder.

3

#12. The event is not very reliable. The **variation in repeated readings of the DV** are considered in relation to the **range of the IV** to determine if a **relationship** can be established.

#13. Is there a large enough change in magnitude of the DV across range of the IV in relation to reliability of the DV to carry on?

#14. **Repeated readings** help assess the reliability. The uncertainty could have been reported and shown on the graph.

#15. The investigator reflects the uncertainty of the data in the claims made about the relationship.

Other people in my class got different results although the pattern was similar – hard-boiled eggs did more spins than fresh ones, but not the same number as mine did. So it depends on the egg as well. This makes it even more difficult to use this method as an egg tester.

Explaining my results

The results fit quite well with what I expected to happen. I still think it is because the runny insides can slosh around inside the shells. When you set it spinning, only the shell and perhaps a bit of the egg spins. The middle must nearly stay still. When it is hard-boiled it will all spin.

#16. There might be variation in the number of spins and time at which the maximum is reached but the pattern is likely to remain the same from egg to egg. Spinning could not be used as an instrument to measure the degree of hardness of an egg (unlike a spring to measure force).

Not all eggs behave the same. The variation in data that would have resulted from using a **large sample** of eggs would have made it harder to establish a pattern.

Questions

1 During the trial run Harbinder said that the spread of results from the uncooked egg might be too big, but after spinning the hard one she thought that she could continue with the investigation. In groups explain why she was concerned.

2 Discuss why Harbinder thought that it would be OK to continue after she'd trialled the hard egg.

3 After trying out the hard-boiled egg Harbinder decided to repeat the egg spinning 30 times for each boiling time. Why did she decide to do this?

4 Harbinder wasn't sure how to draw the graph. Did she get it right? Write notes around the graph commenting on it.

5 Most people don't boil eggs for 30 minutes. Why did Harbinder do this?

6 In the circle of variables, underline the variables that Harbinder controlled.

7 Circle the part of Harbinder's report that comments on the effect of the variables she didn't control for.

8 Why was it important for Harbinder to compare her results with those of others in her class?

9 Draw a star next to the parts of Harbinder's report where she suggests ways to improve her investigation.

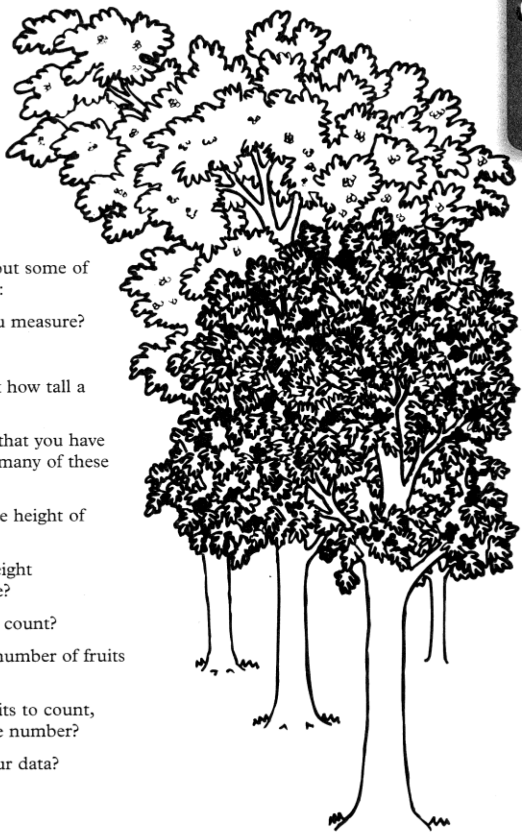
10 Underline information on page 1 that could be used when planning another method for egg testing.

Does size matter?

If you look at trees with fruits on them you'll notice that some trees have more fruits than others. Do taller trees have more fruits than shorter trees?

Find out how the number of fruits on a tree depends on how tall it is.

Complete investigations



You might want to think about some of these things before you start:

- What type of tree will you measure?
- When does it have fruit?
- What factors might affect how tall a tree is?
- How will you make sure that you have reduced the effects of as many of these variables as possible?
- How will you measure the height of each tree?
- How accurate do your height measurements need to be?
- How many trees will you count?
- How will you count the number of fruits on each tree?
- If there are too many fruits to count, how will you estimate the number?
- How will you present your data?

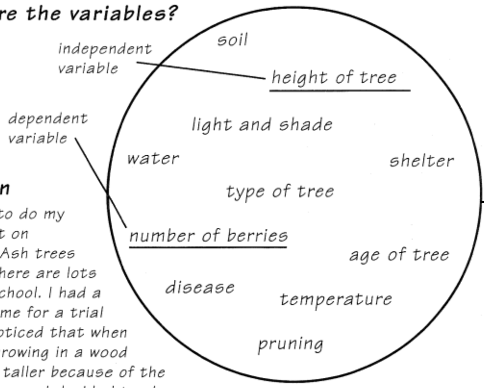
#17. In this investigation the variation is unavoidable. Even when narrowed down to one species of tree there is variation in **sample** of trees; CVs cannot be manipulated; and there is large **uncertainty** due to estimates for **measurements.**

1

Richard's investigation

How does the number of fruits on a tree depend on its height?

What are the variables?



Trial run

I decided to do my experiment on Mountain Ash trees because there are lots near my school. I had a look at some for a trial run and noticed that when they are growing in a wood they grow taller because of the other trees; so I decided to choose trees that were growing on their own.

I will need to survey quite lot of trees because they seem to vary a lot even when they are the same height – there are lots of other things that affect how well a tree is growing, such as soil and how much light there is.

I can't count all the berries. I will count the number in a few bunches that have fallen off the tree and then count the number of bunches on the tree. Since the bunches are different sizes, I'll try to select fallen bunches at random.

I will measure the height of the tree by standing a metre rule against the trunk and stepping back to estimate the height of the tree. I can measure to the nearest fifth of a metre like this, so I think that it is fairly accurate.

What do I think will happen?

I think that the taller the tree, the more berries it will have. But once the tree gets past a certain height it will not have so many berries because it will be old.

My plan

1. *Choose 30 trees that are growing on their own.*
2. *Estimate how high they are using my meter rule.*
3. *Count the number of bunches of berries.*
4. *Find 5 bunches from under different parts of the tree and count the number of berries in each one – use this to get an average for the number of berries for each bunch.*
5. *Work out how many berries there are on each tree.*

2

#18. The trees, whose heights and berries are the focus of this investigation, are affected by many **variables** whose values cannot be manipulated to be kept constant by the investigator.

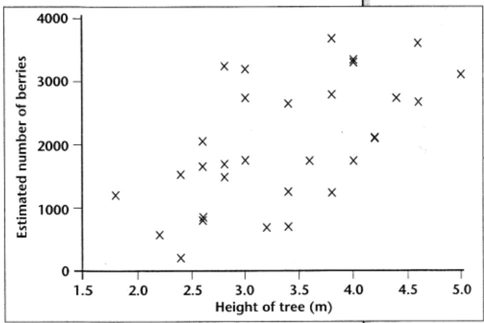
#19. The variation in height due to proximity of other trees could be reduced by **selecting** trees standing on their own. Instead of deciding the values of the IV (as in *Bottle Tip* and *Egg*) the investigator will have to sample enough trees to 'capture' the variation in a **representative sample**. The scale of **uncertainty** in making the **measurements** of 'berries in bunches' and the 'height of the tree' are considered too. The CVs' values could have been measured at each site to enable post hoc **matching**.

My results

Tree	Height (m)	Number of bunches of berries on tree	Average number of berries per bunch	Estimated number of berries on tree
1	2.2	6	94	564
2	4.0	40	83	3320
3	1.8	15	59	1085
4	2.6	12	68	816
5	3.0	31	104	3224
6	5.0	36	86	3096
7	4.2	27	75	2075
8	3.2	10	68	680
9	2.8	17	101	1717
10	3.0	24	74	1776
11	4.6	33	109	3597
12	2.8	37	88	3256
13	3.4	32	84	2688
14	2.4	13	77	221
15	4.4	29	95	2755
16	2.6	32	63	2016
17	2.8	19	78	1482
18	3.0	31	88	2728
19	3.8	45	82	3690
20	3.8	12	103	1236
21	4.0	36	93	3348
22	2.6	18	90	1620
23	4.2	30	70	2100
24	4.6	26	102	2652
25	2.6	9	83	807
26	3.4	18	69	1242
27	4.0	34	52	1768
28	3.6	20	89	1780
29	3.8	29	96	2782
30	2.4	14	110	1540

#20. The variation in much biological data requires large data sets for a pattern to be established.

I made a scatter graph showing the total number of berries and the height of the tree.



#21. The uncertainty in the data is reflected in the data on the scatter plot. Since so many variables might have affected the number of berries (in addition to the height of the tree) a co-relation is all that can be claimed.

Do I believe my results and why?

I think that tall Mountain Ash trees generally have more berries on them than smaller ones. The points are quite scattered from a straight line so the relationship is only a general one. There aren't enough points to be certain how many berries a tree of a particular height will have. I needed to count far more trees to get enough points to be certain about the relationship.

The method of counting the berries in a bunch was quite accurate. The fallen bunches may not have been typical of all the bunches on the tree. I could have counted berries in bunches that were growing. I would need a step ladder to make sure that I didn't just pick those I could reach from the ground and I would have to make sure that the bunches were chosen randomly.

There was a problem counting the number of bunches on the bigger trees. I couldn't remember which ones I'd counted. It might have been better to count them accurately in a small section of the tree and then work out how many on the whole tree. Or I could have repeated the count and found an average number of bunches.

The height of the tree was difficult to estimate, especially with the taller trees. It would have been better to stand exactly 20 m from the tree and measure the angle from the ground to the top of the tree to calculate the height.

#22. An evaluation of the quality of the data draws on the ideas about evidence summarised in the concept map.

Explaining my results

I wasn't surprised that the points didn't fit a straight line. All living things are different so I wouldn't expect the relationship to be very neat. Besides the fact that trees are different, there might also have been differences in the conditions which made them grow or the availability of insects to pollinate the flowers so they could make fruits. I don't know whether the tallest trees were actually the oldest, so my original idea about old trees having fewer fruits cannot be answered by this investigation.

I expect that bigger trees can photosynthesise more than small trees because they have more leaves. The fruits depend on the sugar from photosynthesis to grow. It seems likely that if there's a lot of photosynthesis there'll be more fruit. This is probably why tall trees have more fruit. It would be interesting to find out if they have bigger fruits. There may be other factors that could affect the number of fruits formed.

#23. The explanation draws on the pupils' understanding of substantive ideas; and illustrates how the ideas of evidence (shown with a shadow on the concept map) are informed by the substantive ideas.

Questions

1 Make a table like the one below and list the problems that were identified during the trial run in column 1.

Problems from the trial run	How they were overcome?	Further improvements identified while evaluating
trees seem taller if growing close to others	find trees on their own	none needed
trees seem to vary a lot		

2 Next to each point in your list write down what decision was made to overcome it in column 2.

3 In his evaluation, Richard suggested further improvements. Write these suggestions in the third column.

4 What factors might Richard have taken into account when he decided to measure 30 trees?

5 What is the height of the tallest tree?

4 6 How many berries were on the shortest tree?

7 What data must have been collected but isn't shown in Richard's table of results?

8 Was a scatter graph the best way of presenting the data? Why?

9 None of the control variables could be kept constant. How did Richard make sure that he was carrying out a fair test?

10 What other factors may have affected the number of the fruits formed?