

GCSE MATHEMATICS

HANDLING DATA COURSEWORK

DATA HANDLING COURSEWORK EXAMPLE

Preamble.

The example coursework detailed within these pages is essentially complete and represents what I would say is a typical Intermediate GCSE submission, though one lacking real depth.

It is of course a little fabricated and is intended as a guide for students in writing a report only.

This should probably be read in conjunction with the *Notes on Data Handling coursework* document which gives some advice on the *wheres* and *wherefores* of data handling tasks.

Please note the following:

- i) Important points and terminology are in bold letters. It is very important that students use correct terminology where appropriate.
- ii) This report's final evaluation is quite critical of certain aspects of the investigation including limitations and suggestions for improvements etc. **This is a good source of marks!**
Indeed, in this example, I have to hold my hands up and admit that some of the results did not turn out as expected. **This is NOT a bad thing** providing we discuss possible reasons for the unexpected results and make suggestions for improvements etc.
- iii) This example does not include any statistics necessary for the award of the *higher* marks.
- iv) A hand written report is acceptable, BUT use of I.C.T. is most definitely encouraged!

If I was to mark this example project using published mark schemes, I would probably give the following:

SPECIFY and PLAN:

Appropriate data is chosen. Aims stated in statistical terms (*positive correlation between handspans and reaction distances.*) The sample of people to be measured is explained. Project report is well structured.

Borderline 5/6 marks.

Not 7 marks since practical problems were not foreseen or planned for.

COLLECT, PROCESS and REPRESENT.

Data collected in a form designed to ensure they can be used. Techniques from GCSE grade B descriptors employed with no obvious calculation missing. Very little redundancy in calculation or presentation. Appropriate diagrams used with reasons given for choice of presentation.

6 marks.

Not 7 marks since techniques from GCSE grade A descriptors have not been used.

INTERPRET and DISCUSS.

Summary statistics used to make relevant comparisons and appreciation given that results may not be statistically significant. Allowances made for the nature of the sampling method in making inferences about the population. Effectiveness of the overall strategy has been evaluated and limitations discussed.

6 marks. Almost enough for 7 marks, but ...

Not 7 marks since no plausible reasons given for exceptions. Limited use of probability in the vocabulary used.

TOTAL = 17 marks which is of grade B standard.

I think this is probably on the generous side, particularly as the scope of the project is somewhat limited.

CONTENTS:

Objectives.	3
Sampling.	3
Recording the data.	3
Are reaction distances affected by gender?	4
Are reaction distances affected by handspans?	8
Evaluation.	10

Objectives.

For this piece of work, I will be looking primarily at what I call *reaction distances* of year 11 students. This, **measured to the nearest mm**, is basically the point at which a person catches a ruler between thumb and first finger.

I will investigate the following **hypotheses**:

- do females have smaller reaction distances than males?
- do people with smaller hands have smaller reaction distances?

For the second hypothesis I will be looking for any evidence of **positive correlation** between handspans and reaction distances?

I think that females will have smaller reaction distances because I think that they have smaller hands on average and thus may have more nimble fingers.

To answer these questions, I will need to first collect data from a number of year 11 students. I will then use appropriate statistical methods to analyse my data; including **calculating averages** and using diagrams such as **scatter graphs** to make comparisons etc.

Sampling.

It is obviously impractical to collect measurements from every year 11 student in the country. I will therefore take a **sample** of year 11 students. Such a sample must be **representative** of all year 11 students. So, for example, I cannot simply include only males.

I have decided to select 20 females and 20 males at **random** from my school **population**. I am assuming that there are roughly equal numbers of females and males in year 11 throughout the country.

My school has 58 female and 65 male students in year 11. I used an alphabetical listing and numbered the students from 001 to 123. I then used the *Rnd* button on my calculator to generate sets of random numbers.

The first random number was 0.115. I therefore included the student numbered 115 in my sample.

The second random number was 0.379. I ignored this because there are not 379 year 11 students in the school.

I simply continued like this until I had 20 females and 20 males.

Recording the data.

Having obtained my sample, I then set about recording gender, handspans and reaction distances.

Gender was simply recorded as F for female or M for male. This is known as **qualitative** data.

Handspans were measured to the nearest half cm. They were measured using a ruler from the end of the outstretched thumb to the end of the outstretched little finger. This data is **quantitative** (numerical) and is **continuous** in that the values have been rounded.

Reaction distances were measured by getting the student to open the thumb and forefinger of his/her strong hand about 8 cm. The ruler was held about 2 cm above the outstretched hand and then dropped. At the point where the student caught the ruler was recorded the reaction distance. This was measured to the nearest mm. This data is also quantitative and continuous.

When I finished, I entered all the data into an Excel spreadsheet. This makes it easier to sort the data etc.

The data is shown in appendix A.

Are reaction distances affected by gender?

I will begin by calculating the average reaction distances for both males and females.

The question is which average to use? The **mean, median** or **mode**?

I have decided to use the mean since it is the only average which uses all the data.¹

The mean reaction distance for males is given by the sum of all the distances for males divided by 20.

$$\text{This gives } \bar{M} = \frac{\text{total male distance}}{20} = \frac{278.7}{20} = 13.935 \text{ cm.}$$

Similarly, \bar{F} , the mean reaction distance for females came to 15.975 cm.

So it appears that males have a shorter reaction distance than females.

However, this does not really answer the question since the reaction distances of the males might be much more spread out than the female distances.

To determine this I will draw two **frequency polygons** on the same axes, one for the female distances and one for the male distances. This is a good way of representing **continuous data**.

I will first arrange the reaction distances into convenient groups: 0 to 5cm, 5cm to 10cm etc.

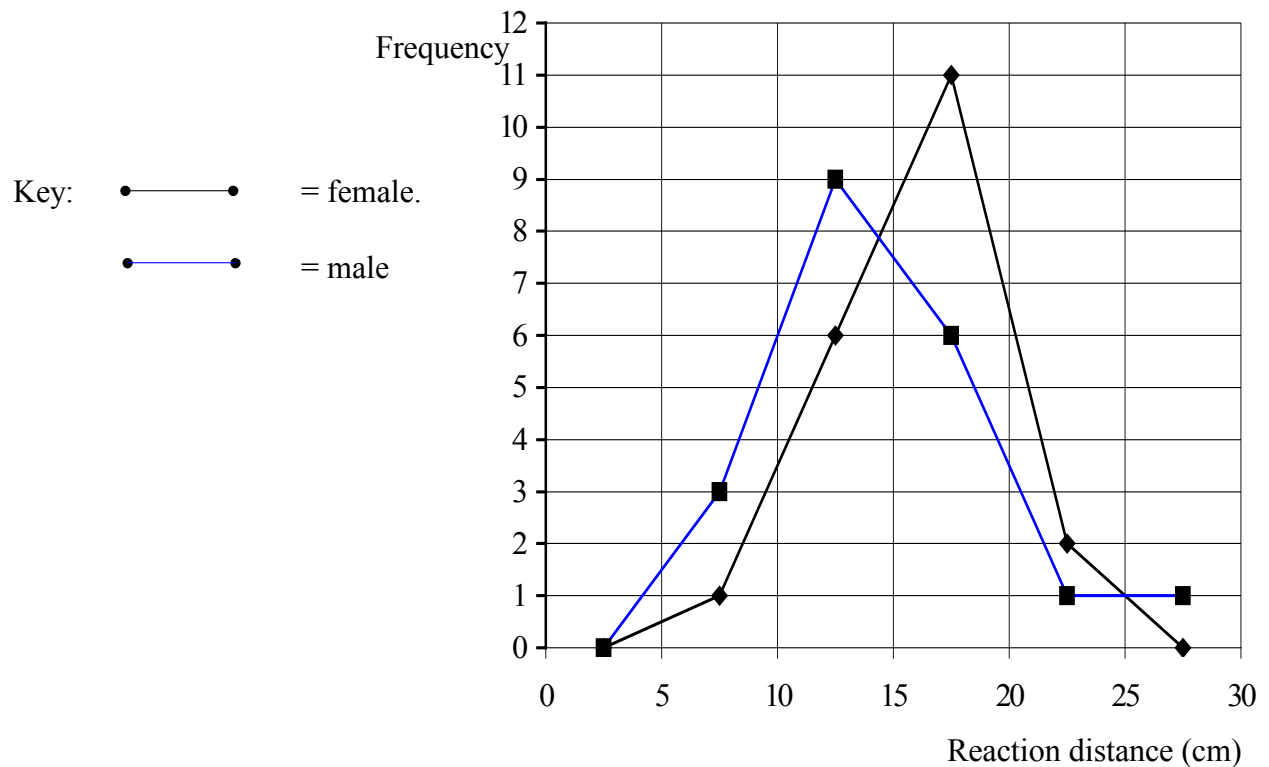
The following **grouped table** shows the results.

	0cm to 5cm	5cm to 10cm	10cm to 15cm	15cm to 20cm	20cm to 25cm	25cm to 30cm
Female	0	1	6	11	2	0
Male	0	3	9	6	1	1

I can now plot the frequency polygons.

¹ You might argue that the **median** is in fact the *best* average to use here since some students might fail to catch the ruler! What do you do if that happens? Obviously, no one failed to catch the ruler in this example!

Frequency polygons of reaction distances.



The graph shows that the mean reaction distance for males is somewhat lower than for females. It also indicates that there is a higher range of male distances than female distances, i.e. the male data is more *spread out*.

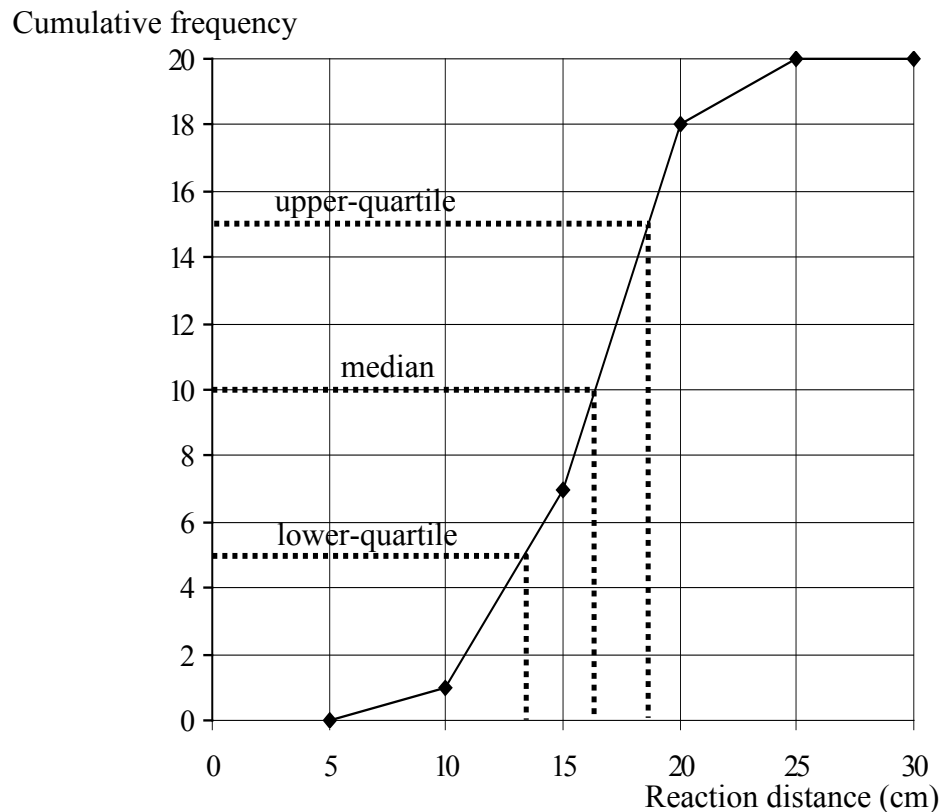
To investigate this further, I will draw some **cumulative frequency diagrams** for the female and male reaction distances in order to determine the **interquartile ranges**.

Cumulative frequency diagram for female reaction distances.

The following table shows the cumulative frequencies for female reaction distances.

Distance (cm)	< 5 cm	< 10 cm	< 15 cm	< 20 cm	< 25 cm	< 30 cm
Cumulative Frequency	0	1	7	18	20	20

I can now plot the cumulative frequency diagram for the female reaction distances.



Using the graph, I obtained the following **statistics** for the female reaction distances:

median reaction distance = 16.5 cm,
lower-quartile = 13 cm,
upper-quartile = 18.5 cm.

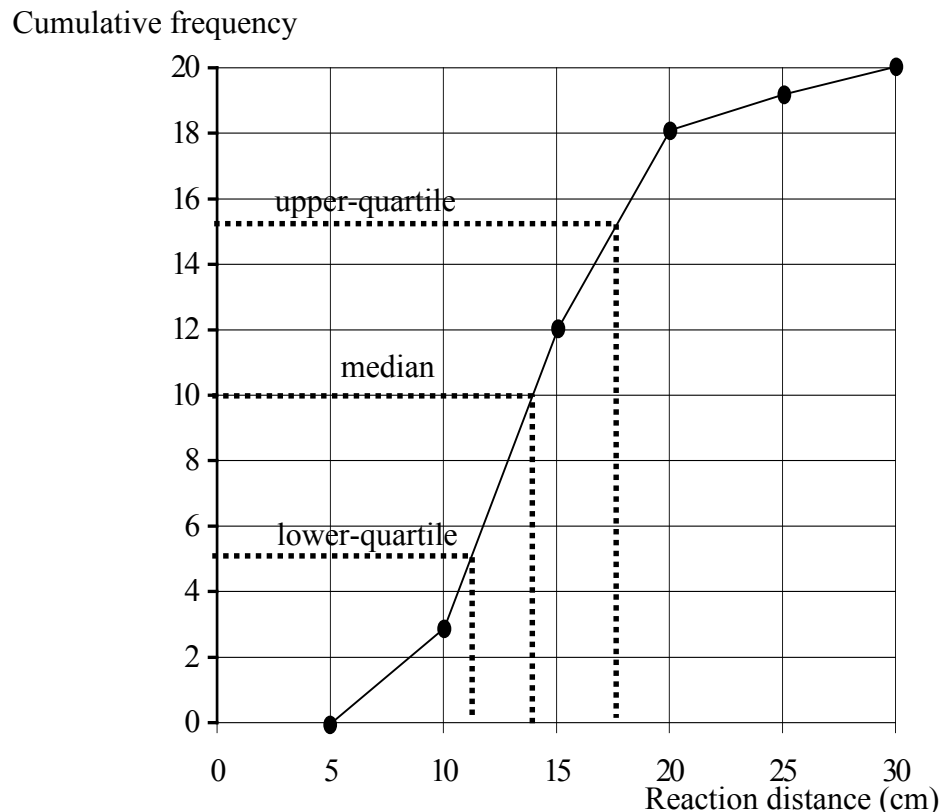
This gives an estimated interquartile range of $18.5 - 13 = 5.5$ cm.

Cumulative frequency diagram for male reaction distances.

The following table shows the cumulative frequencies for male reaction distances.

Distance (cm)	< 5 cm	< 10 cm	< 15 cm	< 20 cm	< 25 cm	< 30 cm
Cumulative Frequency	0	3	12	18	19	20

I can now plot the cumulative frequency diagram for the male reaction distances.



Using the graph, I obtained the following **statistics** for the male reaction distances:

median reaction distance = 14 cm,
lower quartile = 11 cm,
upper quartile = 17.5 cm.

This gives an estimated interquartile range of $17.5 - 11 = 6.5$ cm.

In conclusion then, the slightly larger interquartile range for the male reaction distances means that there was a larger spread of distances than for the females. However, on average, males did have a slightly smaller reaction distance.

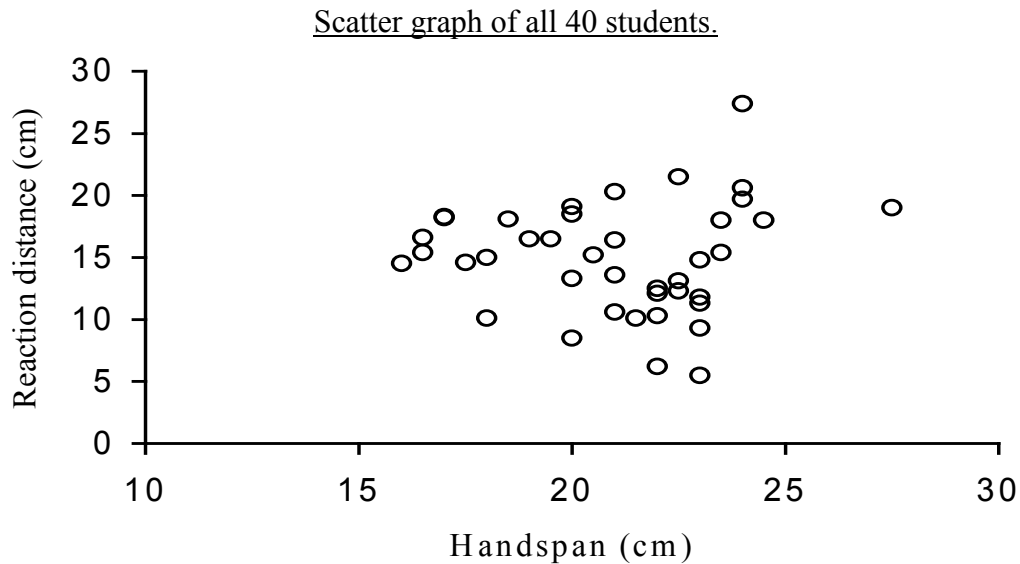
I expected the females to have smaller reaction distances simply because, in general, they have smaller, maybe more nimble fingers etc. The evidence suggests that I am wrong. To be sure, I would need to sample more than 20 of each sex, maybe 100 of each but that would take a lot of time.²

² An example of where the results did not turn out as expected! This is nothing to worry about as long as you give some idea of why this might have occurred etc?

Are reaction distances affected by handspans?

My question in particular is whether a smaller handspan leads to a smaller reaction distance etc? My work above suggests that this may not be true.

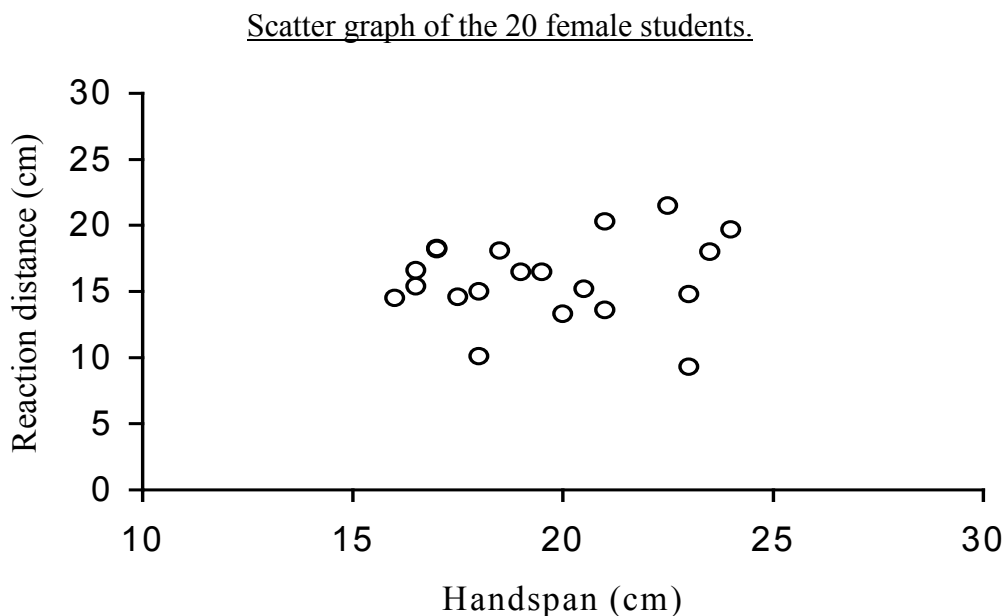
I will begin by drawing a scatter graph of reaction distances against handspans. This should allow me to determine if there is any **correlation** between the 2 sets of data.



In this graph, each plotted point represents a student from my sample.

The scatter graph indicates only a **weak positive correlation** between handspans and reaction distances. I therefore think it inappropriate to draw a line of best fit in order to estimate reaction distances.

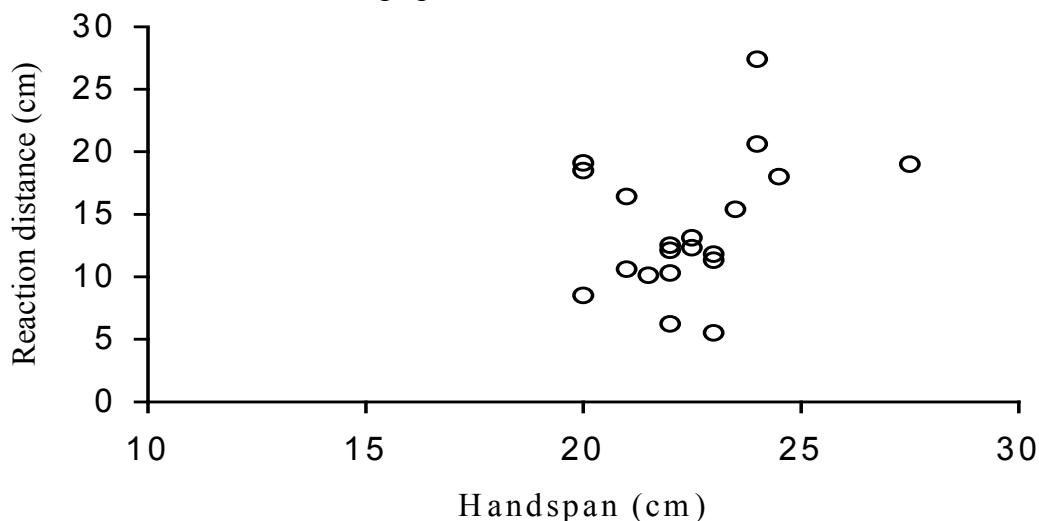
The next scatter graph shows only the female students.



Again, there is very little correlation.

Finally, I will plot a scatter graph for the male students only.

Scatter graph of the 20 male students.



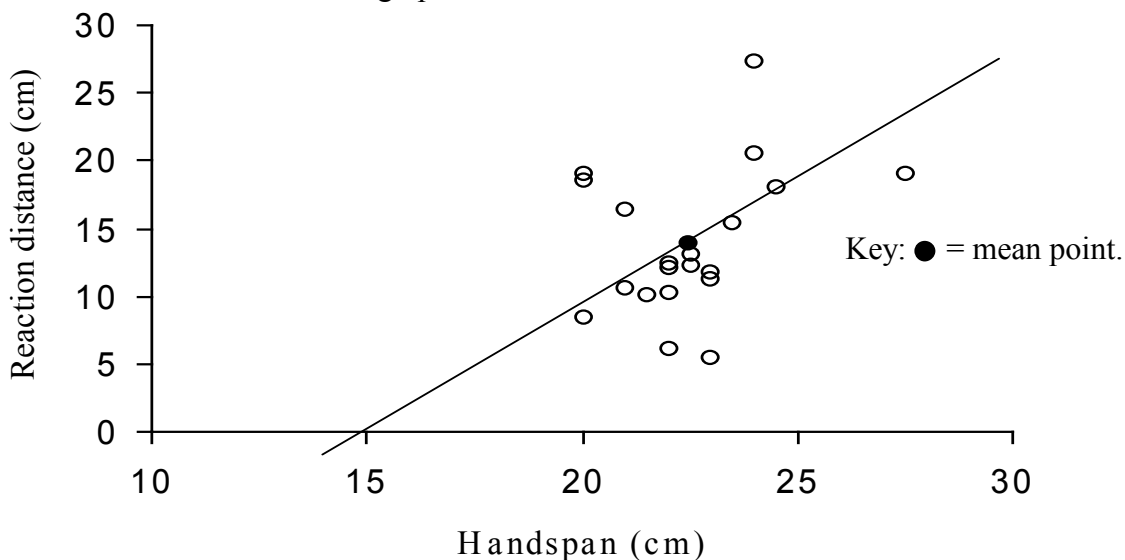
Now this time, there is evidence of **positive correlation** between handspans and reaction distances for the male students only.

This enables me to plot a line of best fit on the above scatter graph for the male students.

My line should pass through the mean point. Now the mean male reaction distance was found to be 13.935 cm.

The mean MALE hand span is given by $\frac{\text{total male handspan}}{20} = \frac{449}{20} = 22.45$ cm.

Scatter graph of the 20 male students.



So, for example, using my line of best fit, I would estimate that a year 11 male student with a handspan of 20 cm should have a reaction distance of about 10 cm.

To answer the question stated above of whether a smaller handspan leads to a smaller reaction distance, I would say that the evidence suggests that this may be true for year 11 male students, but not necessarily true for female students.

If I am correct in this conclusion then it means that male reaction distances are more predictable than female distances.

Evaluation.

To conclude this investigation I will produce a little evaluation of my results.

Firstly, since my analysis and results were based upon a random sample of only 40 students, I have to say that my results cannot be guaranteed to be correct. Given the time, I would have preferred a sample size of say 100 or even 200. Also, for practical reasons my sample was obtained from the students in my own school and not from the entire **population** of year 11 students in the U.K. This could mean that my results are **biased**. In fact, it could be possible that a sample from a different school would produce completely different results.

I am also aware that my recording of the reaction distances was probably not entirely accurate in that it was difficult to judge how high to hold the ruler above the students outstretched hand etc.

Having said this, I am quite happy with the results obtained from my data. For example, I feel that I am justified in saying that a typical year 11 female student from my school is likely to have a reaction distance between 13 cm and 18.5 cm (between the lower and upper quartiles) and that the corresponding range for males is between 11 cm and 17.5 cm. This gives a nice illustration of the differences I found between female and male reaction distances.

When looking for any links between handspans and reaction distances, I did not find the kind of correlation I was expecting. I expected there to be positive correlation for both females and males.

However, there was certainly evidence of positive correlation in the male data which I used to construct a line of best fit etc. This must be used with care however. For example, the line predicts that a male year 11 student with a handspan of 15 cm will have a zero reaction distance! This is clearly ridiculous and suggests that there is actually very little correlation between handspans and reaction distances, that again my sample may be too small. The point is that the line of best fit should not be used for handspans less than say 20 cm or more than 30 cm.

I think that my work suggests that there is actually very little correlation between handspans and reaction distances; that the size of a persons hand does not necessarily affect that persons reaction distance. To be sure, a much larger sample would be required.

It might be worthwhile extending the investigation to see if reaction distances decreased with practice; e.g. compare a persons initial reaction distance with that obtained on say the 5th attempt. To see if females improved as much as males etc. Maybe an investigation into links between reaction distances and ability to estimate angles and distances would prove useful.

APPENDIX A

Pupil	Gender	Handspan (cm)	Reaction distance (cm)
115	M	27.5	19.0
007	F	17	18.2
036	M	23	5.5
091	F	23.5	18.0
120	F	16	14.5
002	M	22.5	13.1
047	M	24	27.4
053	F	17	18.3
013	M	21	10.6
086	M	20	8.5
054	M	22	10.3
016	F	21	13.6
109	M	24.5	18.0
021	M	24	20.6
112	M	20	18.5
014	F	18	10.1
031	F	18	15.0
071	M	22	6.2
065	M	22	12.5
015	F	17.5	14.6
032	M	21.5	10.1
026	F	21	20.3
088	M	23	11.3
119	M	21	16.4
079	F	23	9.3
068	F	20	13.3
017	F	20.5	15.2
104	M	23	11.8
095	F	18.5	18.1
048	F	24	19.7
023	F	16.5	15.4
006	M	22.5	12.3
019	F	23	14.8
028	F	16.5	16.6
033	F	19	16.5
098	M	20	19.1
004	F	22.5	21.5
061	M	23.5	15.4
025	F	19.5	16.5
005	M	22	12.1