National Certificate of Educational Achievement
TAUMATA MĀTAURANGA Ā-MOTU KUA TAEA

Exemplar for Internal Assessment Resource

Mathematics Level 3

Resource title: Estimation

This exemplar supports assessment against:

Achievement Standard 91583

Conduct an experiment to investigate a situation using experimental design principles

Student and grade boundary specific exemplar The material has been gathered from student material specific to an A or B assessment resource.

Date version published by Ministry of Education	December 2012 To support internal assessment from 2013

For this task students are required to conduct an experiment to investigate a situation using experimental design principles and produce a report produce a report describing an experiment investigating a factor that influences the accuracy with which people estimate half of a continuous quantity.

	Grade Boundary: Low Excellence
1.	For Excellence the student is required to conduct an experiment to investigate a situation using experimental design principles, with statistical insight. This involves integrating statistical and contextual knowledge throughout the investigation process, and may include reflecting about the process; discussing how possible sources of variation were dealt with during the design phase; considering other relevant variables. There is evidence of: • Background contextual knowledge by referencing research related to using anchors to influence estimates when people are uncertain, leading to an appropriate question (1) • A logical and clearly described plan for the experiment (2) • An appropriate formal statistical inference by assessing and interpreting the strength of evidence about the causal relationship (3) • Consideration of broadening the experimental situation in the discussion of the student's findings (4). For a more secure Excellence the research needs to be more clearly linked to the experiment. For example, the student could develop the comments on why 250 mL was a good choice for the volume of liquid inside the bottles, and how the results might change if the bottles were changed.

Student 1: Low Excellence

For my experiment I want to determine if the estimates of how much water is in a bottle can be affected by the size and shape of the bottle. My experiment therefore is based on estimating how much water is in a bottle. I want to find out if you could get people to give higher estimates then the real quantity by using a 600mL bottle than you would get when people look at a 1.5L bottle. I spent some time researching the possible results for each bottle size. In my research, I found that 250mL was a good basis for the experiment as it allows both the 600mL bottle and 1.5L bottle to have a reasonable amount of liquid in without showing the student exact amount. For example if I had used 300mL then the 600mL bottle would have been half full and this may have been easy for the students to guess, secondly the 1.5L bottle would be one fifth full and again this may have been easily guessed whereas 250mL of liquid would reduce the chance of students being able to guess the correct amount. I found from my research also that using 250mL students were more likely to over-estimate the volume in 600mL bottle and are more likely to underestimate the volume in a 1.5L bottle. The research about bottle size also showed that when people are uncertain about the volume they are more likely to over-estimate. The research suggested that the smaller the bottle, the higher the estimate would be. I think that a larger bottle will make the amount of water look smaller, because it will be a smaller proportion of the bottle filled than a smaller bottle. (1)

In order to conduct this experiment, I will use 53 students from Year 12. There will be two independent groups for this experiment. One group will be shown and asked to estimate the amount of liquid in the smaller bottle (250mL) and the other group will be asked to estimate the volume of liquid in the bigger bottle (250mL).

The treatment variable for this experiment will be the size of the two bottles. One of the bottles will be big (1.5L), and one will be small (600mL) but both will contain the same amount of water (250mL). I thought about using bottles that were not familiar sizes as this would not allow such an easy estimate to be obtained e.g. half full of the 600mL bottle, so around 300mL. However, upon reflection I decided that using the familiar coke bottle 600mL and 1.5L would give better and more consistent estimate than using non familiar bottles such as dish washing liquid bottles.

One of the groups will be asked to estimate how much water there is in the small bottle, and the other will be asked to estimate how much water there is in the larger bottle. This will give us the answer to our experiment which is too see if the size of a bottle will change a student's perception of the amount of water that there is in the bottle. If we had given both groups the same sized bottle then we would not have been able to compare the results to see if the size of the bottle made a difference, because there would be no other different sized bottle to compare the results to. The response variable will be the estimation that is made by students. This estimation will be based on the volume of water that each treatment group sees in the bottle to the nearest millilitre.

To minimise bias for this experiment we will assign one independent person to take the larger bottle containing 250mL to one of the groups, and another person to take the smaller bottle of water containing 250 mL to the other group. This is done so to avoid the students from seeing the two bottles together at any time. If they saw the two bottles together then they may get confused or, compare the amount of water in both bottles and they may guess that they two bottles share the same amount of water in them if they see the two bottles together, and so ruin the experiment because the main goal is to keep the students from knowing that the two bottles contain the same amount of water. (2)

<u>Variables that can be controlled:</u>

The type of class chosen for our experimental units:

To minimise other factors for this experiment I am going use two maths classes, instead of for example using one maths class and a P.E class. I did this because it is more likely than not that the students of the P.E class will carry drink bottles of the same kind that is used for this experiment in comparison to the students of the maths class. I then randomised each group. To do this I selected two year 12 mathematics classes that occurred at the same time of the day. This gave me a total of 53 students. I then went into the first class and asked each student to line up outside as they were doing this I went into the next class and asked each student there to also line up outside with the students from the first class. I then asked each student to select a piece of paper. On the piece of paper there was either a 1 or a 2. Each piece of paper was of identical size and shape. Once every student had a piece of paper I then asked that all the students with one's to go to room 32 and all the students with two's to go room 34. The group which had the piece of paper with the 1 on it were then asked to estimate the volume of liquid in the 600mL bottle and the group in room 34 with the piece of paper with the two on it were asked to estimate the volume of liquid in the 1.5L bottle. Both bottles contained 250mL of liquid. (2)

Whether or not there were any water bottles on sight, expect for the bottle used for the experiment:

I ensured that all students were asked to put all their belongings away from their desk as the beginning stage of conducting the experiment. Instead of asking the students to remove their drink bottles, I made it more general saying everything off their desk, so that they would not suspect that the experiment would be about their water bottles, so that it would not affect their estimations. If I had told them to specifically remove only drink bottles from their table, then there was a high chance that they may have clicked that the experiment had something to do with water bottles and analysed their drink bottle before putting it away. This in turn would give them an advantage when it came to estimating the amount of water there was in the bottle because they may remember how big their own bottles were, and if they had the same bottle then they would know the size of the bottle used for the experiment.

The location of the experiment:

I made sure we did not conduct the experiment on a P.E class or any other class that may be on the field. A class that was on field would be much harder to control. I needed to have a high amount of control over the students so that they would not be able to discuss their answers amongst each other, giving some students an advantage in estimating and not finding out the true estimate of each individual. This would not have been the case if they were able to talk amongst each other.

The time that the experiment was conducted:

I ensured that I did not conduct the experiment close to the time the bell times, for example the end of a period, or the start of interval or lunch. I did this because if we did not finish the experiment in the time, the students will have changed their classes and this could mean any of the following: that they would be able to talk to each other between classes and discuss answers and had I chosen classes of different subjects then the same students may have been included in the experiment again. To ensure that this did not happen, I made sure I arrived at the classes earlier on in the period, very much towards the start and then went to the second class immediately after the completion of the experiment. This ensured that both experiments were conducted at approximately the same time of day, which allowed us to make sure no one from either of the class were able to communicate. If either of the groups had been able to communicate with then they might have realised that the two bottles contained the same amount of water ruining the experiment.

Whether or not the bottles had labels on them:

I ensured that the bottles did not have any labels on them so that the students would not be able to see the amount of water that the bottle initially is able to carry. I did this so that they did not have anything to go on when estimating the amount of liquid in the bottle. I did this to eliminate any advantage that they could have had while they were estimating the amount of water in the bottle.

Same information being told to each group:

For the same information being told to each group, this will include students being told what to do. Neither group will be told the size of the bottles or the amount of water in each bottle. An example of a variable that cannot be controlled is the amount of bottled water that each student consumes. This makes it easier for them to correctly guess the amount of water in the bottle if they recognize it as a brand they are commonly exposed to. This variable is unable to be avoided, as I have no knowledge of how good a student's ability to guess the size of bottles is.

The experiment will be carried out over the space of one normal school period (55 minutes).

I will visit one group at the first half of a period, and I will visit the other group immediately after to ensure the students are guessing at the same time of day and to also ensure that they do not come into contact with each other over the duration of the experiment. Each group will be shown one of the two bottles, each one containing the same amount of water (250mL), then they will be asked to estimate how much water they think is inside the bottle in millilitres. I recorded the information before I left each group to ensure the recorded estimates were not tampered with or altered following the initial estimation.

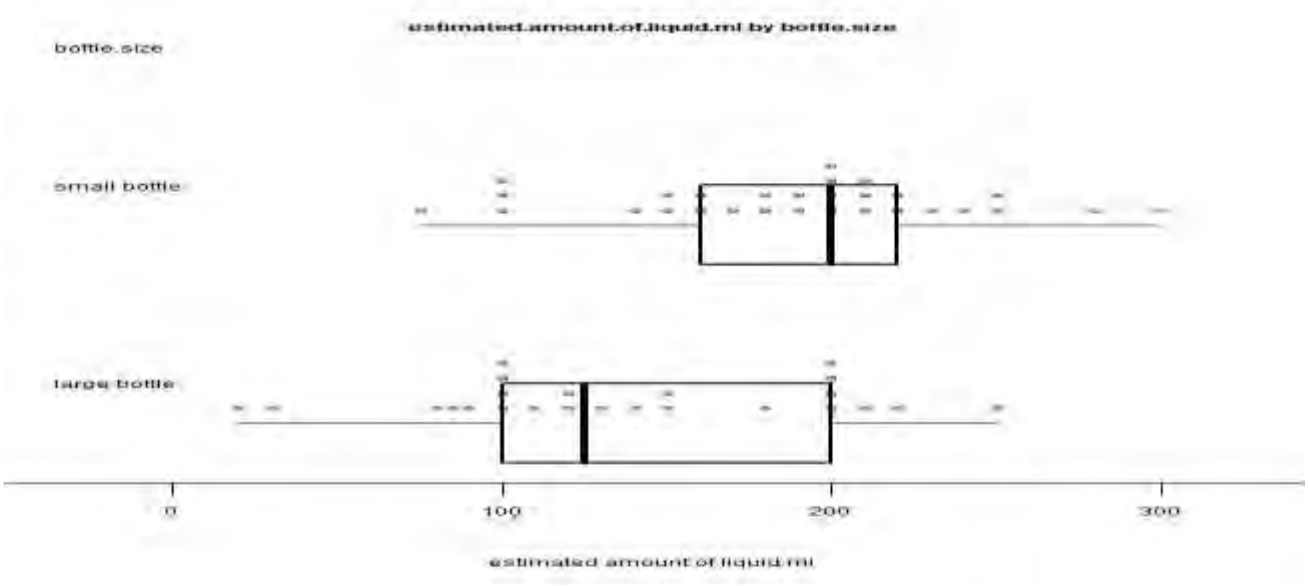
<u>Variables that cannot be controlled:</u>

A Students natural ability to estimate:

I was unable to separate the students who were naturally more skilful at estimating from the ones which weren't.

Gender:

I was unable to make sure there was a same amount of female and male experimental units between the two classes.



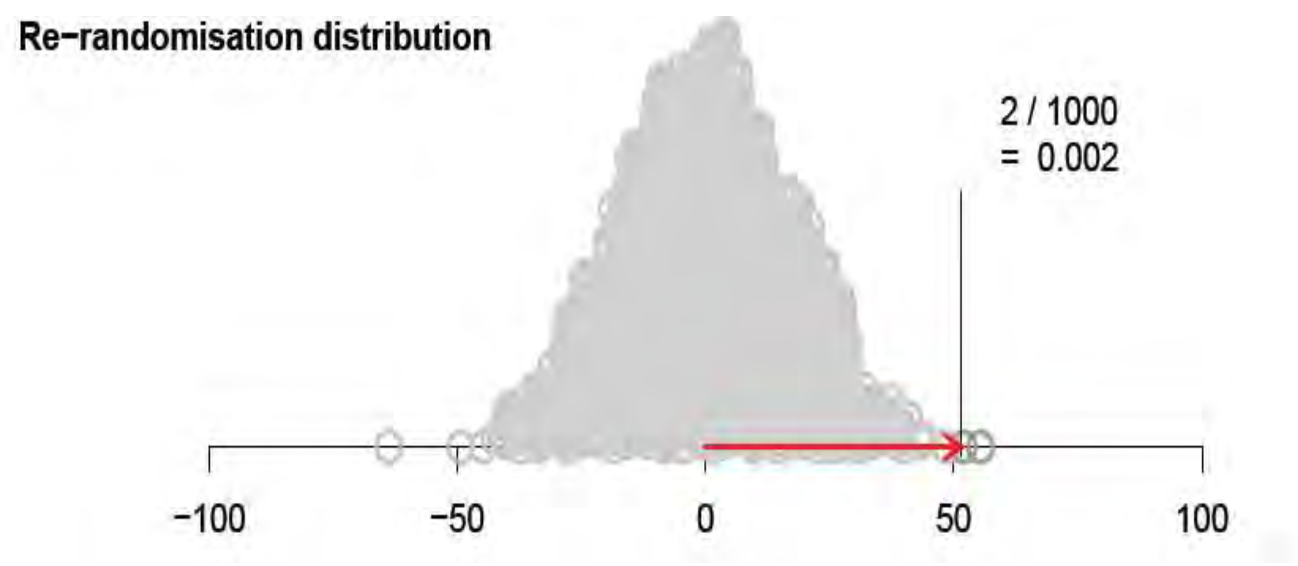
Summary of estimated.amount.of.liquid.mL by bottle.size

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	Sample.Size
large bottle 1.5L	20	100	125	136.9	200	250	60.141	26
small bottle 600mL	75	160	200	188.4	220	300	53.604	27

The first thing I noticed was that both of the groups had medians and means that were less than the actual amount of water in the bottles, and the median amount for the smaller bottle group was 50mL away from the actual 250mL. Nearly everyone in the experiment estimated 250mL or below, for the amount of water in the bottles. The two very low estimates for the big bottle I considered removing because I think they were being silly (30mL and 20mL) but maybe they really do not know how to estimate volume.

There is a difference of 51.5mL between the two group's means. This difference, suggests that the 600mL bottle estimates while higher than that of the 1.5L bottle estimates is still not sufficient for me to say that people looking at a 600mL tend to guess higher than those looking at a 1.5L bottle.

Also, the variability in these estimates for the 600mL bottle is much smaller than the 1.5L bottle – you can see this in the box-and-whisker plot, as the box for the 1.5L bottle is much wider than that of the 600mL bottle. This shows that people who were in the 1.5L group were less consistent in their estimates, and most gave values quite a bit lower than the 250mL of liquid. Students who were in the 600mL group gave estimates a lot closer to the actual amount as shown with a median of 200mL and their estimates are roughly symmetrically distributed around the median of 200mL. I have decided to focus on the difference between the means of the two groups (I could have also used the medians) to look for evidence to answer my investigative question. To explore whether it is highly unlikely to get a difference as big or bigger than my observed difference of 51.5mL by chance alone, I randomly re-assigned the estimates to the two groups. Each observed value was randomly re-assigned (re-randomised) to one of the two groups and the difference between the re-randomised group means was calculated. This was repeated 1000 times producing, under chance alone, 1000 differences between the group medians. The graphs and results produced from this method are shown below:



From the re-randomisation distribution, I see that the difference I got (51.5mL), or greater, happened 2 times out of 1000. That is a probability of 0.02%.

As the estimates produced by random allocation of 0.02% are at least as far from zero as the observed estimate, then the data provides very strong evidence of a link between the two variables. This means that because the probability is very low, it would be very unlikely that a difference of 51.5mL could happen by chance, so something else must be working with chance to explain the effect.

The results from the randomisation test, and the fact that my experiment was well- designed and executed, means that I can claim that students do give higher answers by using a 600mL bottle than you would get when people look at a 1.5L bottle for the same volume of water in each bottle. My results are also important in terms of how much liquid to place in a bottle as too little or too much could conceivable allow the student to estimate the amount of liquid more easily whereas 250mL of liquid provided a good basis for people to estimate on. (3)

I wanted to find out if you could get people to give higher answers by using a 600mL bottle than you would get when people look at a 1.5L bottle. If I had a chance to do this experiment again I would use different bottles of different shapes rather than use the common type and I would choose maybe a dish washing bottle and a detergent bottle for students to estimate, this could minimise the influence of personal knowledge about how full a bottle looks.

Another point is when they answering the question I will give every student only 2 minutes to control another variable of the time use to answer. I would also show each student the bottle with liquid in it separately. This would ensure no contact between students when they were estimating the amount of liquid present in the bottle. These improvements could make some difference to the estimations of the students because they have to estimate how much liquid is in the bottle within a controlled time and area. This could improve the experiment to have a statistical significant and getting a better result.

From the reading I did, I learnt that boys are better at estimating than girls. It would be interesting to see if I randomised the classes by gender only and kept the amount liquid the same in one bottle i.e. 600mL if these results could be replicated by our group. (4)

Appendix

Results:

BOTTLE SIZE	ESTIMATED AMOUNT OF LIQUID
large bottle 1.5L	200
large bottle 1.5L	100
large bottle 1.5L	85
large bottle 1.5L	200
large bottle 1.5L	30
large bottle 1.5L	20
large bottle 1.5L	250
large bottle 1.5L	100
large bottle 1.5L	120
large bottle 1.5L	100
large bottle 1.5L	80
large bottle 1.5L	180
large bottle 1.5L	100
large bottle 1.5L	200
large bottle 1.5L	150
large bottle 1.5L	140
large bottle 1.5L	120
large bottle 1.5L	90
large bottle 1.5L	110
large bottle 1.5L	130
large bottle 1.5L	150
large bottle 1.5L	200
large bottle 1.5L	210
large bottle 1.5L	220
large bottle 1.5L	100
large bottle 1.5L	85

BOTTLE SIZE	ESTIMATED AMOUNT OF LIQUID
small bottle 600mL	250
small bottle 600mL	100
small bottle 600mL	100
small bottle 600mL	200
small bottle 600mL	300
small bottle 600mL	200
small bottle 600mL	200
small bottle 600mL	200
small bottle 600mL	250
small bottle 600mL	280
small bottle 600mL	100
small bottle 600mL	75
small bottle 600mL	240
small bottle 600mL	220
small bottle 600mL	210
small bottle 600mL	220
small bottle 600mL	210
small bottle 600mL	190
small bottle 600mL	180
small bottle 600mL	170
small bottle 600mL	190
small bottle 600mL	210
small bottle 600mL	230
small bottle 600mL	160
small bottle 600mL	150
small bottle 600mL	140
small bottle 600mL	150

	Grade Boundary: High Merit
2.	For Merit the student is required conduct an experiment to investigate a situation using experimental design principles, with justification. This involves linking components of the process of investigating a situation by experiment to the context, explaining relevant considerations in the investigation process, and supporting findings with statements which refer to evidence gained from the experiment. There is evidence of: • An appropriate causal relationship question with a prediction of likely outcomes (1) • A description of the type of experiment, identification of the variables and how the treatment was allocated to the experimental units (2) • A formal inference by assessing the strength of evidence for the causal relationship (3). To reach Excellence the student would need to provide evidence of investigating a situation, with justification, by more clearly linking the discussion to the context.

Student 2: High Merit

My experiment is about guessing the age of a person by looking at the picture of a person. I will investigate whether wearing make-up has an effect on how people guess the age. In my research I have found that celebrities wear makeup aimed at making them look very much younger than their real age and making it harder to guess their age. I have found that celebrities are hardly seen without makeup and if they are seen without makeup the media goes crazy. From the information/ideas found from research for this experiment, it shows that comparing two photos, one with makeup on and another one without make up, can vary the estimated age of a person. The photos of a person with makeup and another without make up can lead people toward an answer which is different, because makeup affects the look of a person by covering dark and red spots, wrinkle, discoloration area, breakouts and any other undesirable spots or areas on their face. These things make photos of a person with and without makeup guide people into different estimates of the real age. I'm not sure if wearing makeup will always make people look younger, because often people wear makeup to look older (like teenagers), so my problem for this investigation is "Does changing the picture of a person wearing makeup and without makeup have an effect on the guesses of the celebrity's real age?" (1)

I will investigate this problem by using Kim Kardashian photos that will be found from Internet for student to estimate; by the way her real age is 32. There are two treatment groups in this experiment: one is the photo of Kim with her makeup on and another one is the photo of Kim without her makeup, to see if there is a difference of the estimation of her age between the two photos/groups. These photos have a question "How old is she? " underneath.

How old is she? How old is she?

They will be printed in black white and cut individually. There are 50 photos of Kim altogether, 26 identical photos with makeup and 24 identical photos without makeup. The experimental group will be year 13 students a total of 50 people. I have to ensure that the 50 students selected for this experiment both have a maths class at the same. The first class had 26 students and the second class had 24. I went into both classes one immediately after the other to ensure that student did not have an opportunity to discuss their results. For the class of 26 students I gave each student a black and white photo of Kim Kardashian with make-up on, and asked them to answer the question beneath it. When they had finished I asked them to raise their hand and I collected the photos with the answer to the question back. I then went into the second class and repeated the process with one exception that the photo this group saw was one of Kim Kardashian without make-up. There will be no time limit. I will record the data from the collected answers on to a spreadsheet. The response variable of this experiment will be the estimated age of Kim Kardashian in years. (2)

The variables we can control in this experiment are

- Each group will answer the question in the same test condition, same date, same time and are in the same group.

- The students will get the same photos of Kim Kardashian, which is separated into two treatment groups of with makeup and without makeup.

- The photos that the students will get, will be printed in black and white with the same question "How old is she?" underneath.

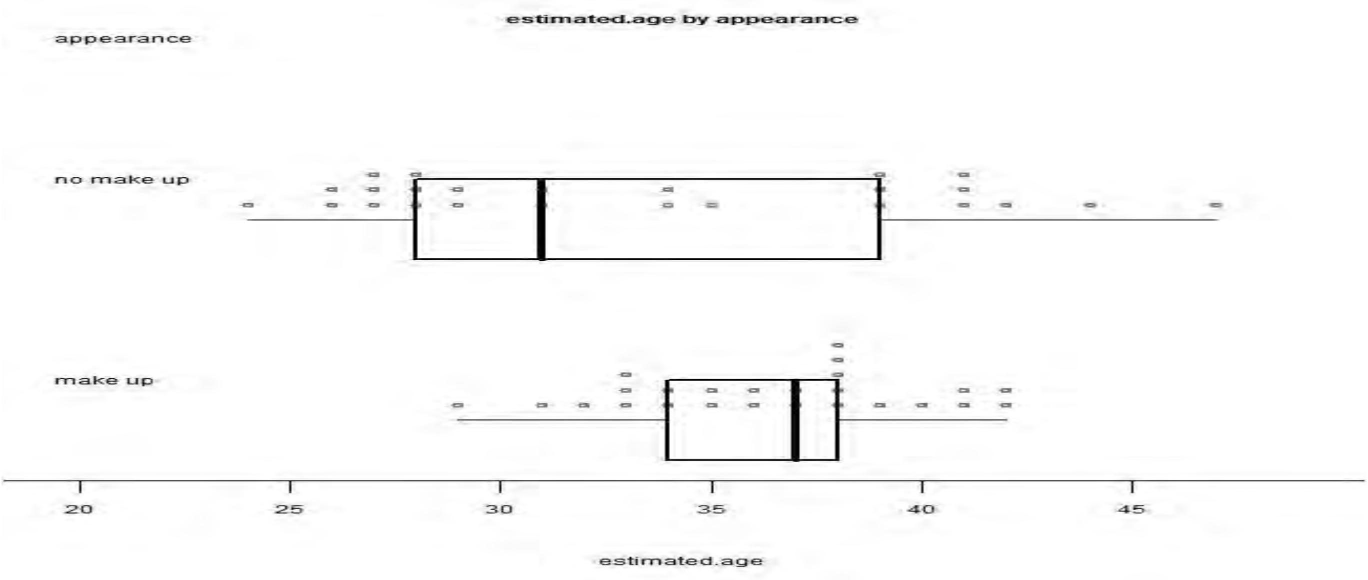
- The students will get the photos at the same time and the photos will be collected in at the same time.

- I will hand out the photos to the students and I will be the person who records the data.

The variables that we cannot control are:

- The personal knowledge of the students because some students might know the age of Kim Kardashian from magazine, Internet or TV shows.

- Some of them might just guess random answers



Summary of estimated age by appearance

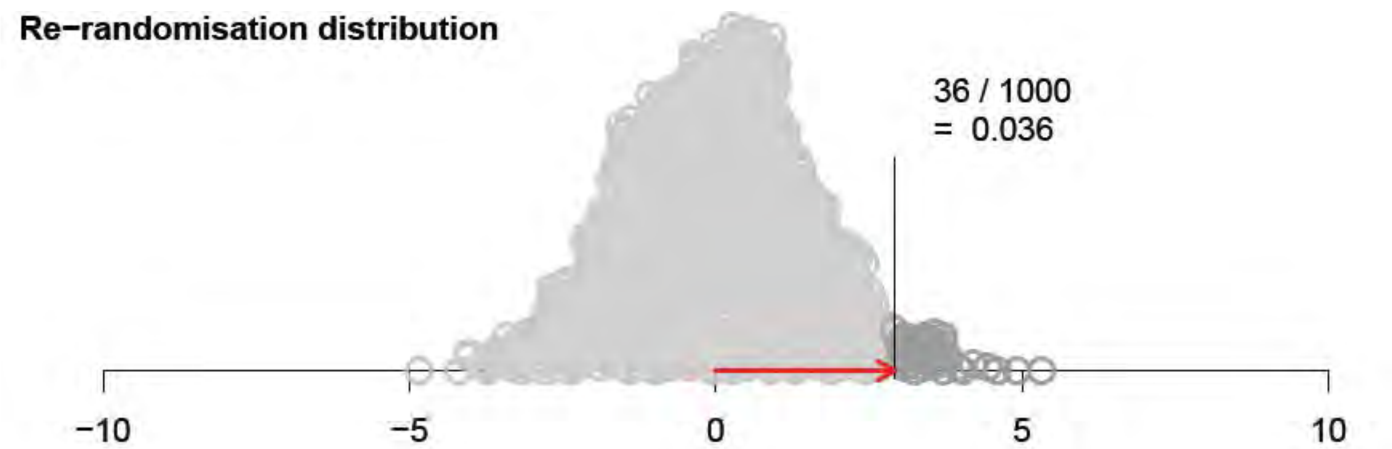
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	Sample.Size
make-up	29	34	37	36.40	38	42	3.488	25
no make-up	24	28	31	33.48	39	47	6.771	25

The box and whisker plot of the data shows that the difference of the two medians is 6, whereas the difference of the means of the two groups is 2.92 years. The distributions for both make-up and no make-up are pretty close to being symmetrical.

The estimations of the make-up group are pretty close together while the no makeup group is spread out, which means that their estimations are more varied. The standard deviation for the no make-up is higher – this combined with the larger spread shows me that people were less sure of the age of Kim wearing no make-up when they were shown the photo.

When I compare the features of the graphs for the two groups, they do not look like what I would expect if chance was acting alone.

I used the difference between the means of the two groups to look for evidence to answer my investigative question. I need to find out if it is likely to get a difference as big or bigger than 2.92 by chance alone. I used the randomisation test 1000 times to produce just by chance 1000 differences between the group means. The graph and results produced from this method are shown below:



My observed difference of 2.92 years only came up 36 times out of 1000.

As the estimates produced by random allocation of 3.6% are at least as far from zero as the observed estimate, then the data provide some evidence of a link between the two variables. This means that because the probability is low, it would be unlikely that a difference of 2.92 years could happen by chance alone, so something else must be working with chance to explain the effect. This might happen because the experiment is too weak; the photos that we used in the experiment might be too similar and did not make much difference to the student's estimations.

The results from the randomisation test, and the fact that my experiment was well- designed and executed, means that I can claim that students do not give higher answers by changing the picture of a person wearing make-up and without make-up have an effect on the guesses of the celebrity's real age. My results are also important in terms of how old make-up can make you look. (3)

If I had a chance to do this experiment again I would use photos of an ordinary person when wearing make-up and without wearing make-up for students to estimate, to control the variable of personal knowledge because using Kim Kardashian who is a celebrity, most of people might already have some knowledge about her, making this knowledge a variable that we cannot control. So therefore using someone so well known such as Kim Kardashian, where teenagers are more than likely to know her actual age may have an affect on the outcome of the experiment. Something else I would do would be to give each photo in colour. Another point is when they answered the question I would give every student only 2 minutes to answer. These improvements would make some difference to the estimations of the students because they have to estimate age of a person who they never know before within a controlled time. The results could be affected by the type of makeup the person used in the picture, for example if the girl in the picture used less makeup and make it more subtle she would have looked younger. The time the photo was taken could have affected the student's answers and also their gender. Boys probably don't know much about makeup and just wrote a random age. Girls would have looked longer and seen that she's wearing makeup and maybe put her age as lower.

Appendix

Results:

APPERANCE	AGE
make up	37
make up	35
make up	38
make up	38
make up	36
make up	33
make up	41
make up	41
make up	36
make up	33
make up	34
make up	40
make up	38
make up	33
make up	39
make up	31
make up	32
make up	37
make up	38
make up	35
make up	42
make up	42
make up	29
make up	38

APPERANCE	AGE
no make up	34
no make up	39
no make up	39
no make up	27
no make up	44
no make up	41
no make up	41
no make up	47
no make up	29
no make up	41
no make up	28
no make up	28
no make up	39
no make up	27
no make up	26
no make up	31
no make up	35
no make up	28
no make up	42
no make up	24
no make up	26
no make up	29
no make up	27
no make up	34
no make up	31
no make up	34

	Grade Boundary: Low Merit
3.	For Merit the student is required conduct an experiment to investigate a situation using experimental design principles, with justification. This involves linking components of the process of investigating a situation by experiment to the context, explaining relevant considerations in the investigation process, and supporting findings with statements which refer to evidence gained from the experiment. There is evidence of conducting an experiment to investigate a situation using experimental design principles, with justification. • The plan justifies the definition of the treatment and response variables (1) • The student links the features of the data to the features of the experiment and the need to do the randomisation test (2) • The student correctly interprets the results of the randomisation test which is used to assess the strength of evidence (3). For a more secure Merit the student would need to provide more detailed explanations of some of the considerations in the investigations. For example the treatment and response variables are not fully justified and other sources of variation have not been fully described.

Student 3: Low Merit

I wonder if having some objects (dots) spread out will affect the number of objects (dots) estimated? My experiment was about estimating the number of objects (dots) and whether the estimates could be affected by how close the objects (dots) were displayed. I wasn't actually sure what I thought I would find – on one hand I thought that maybe the dots displayed close together might give higher estimates because you can see them with one look and find groups of ten easier, whereas if they are displayed far apart it would harder to see them all in one look so you might not find that many. But on the other hand, maybe having the dots spread out covering the whole page makes it look like heaps of dots. (1)

I carried out an experiment to compare the estimates from dots on a page bunched together and a page of dots that are spread out. I used two independent groups for our experiment so people were not in the same group and only estimated the number of dots for one situation (close together or far apart).

In class, I worked on Microsoft PowerPoint to design my slides that I was going to show to the class. For one group, the first slide I had was blank, so that it would give the students time to get ready. On the second slide I had 20 dots bunched very close to each other so they were touching but not making patterns that could be recognised. This slide was 3 seconds long. The third slide I wrote on it that they only had 5 seconds to write down their estimates.

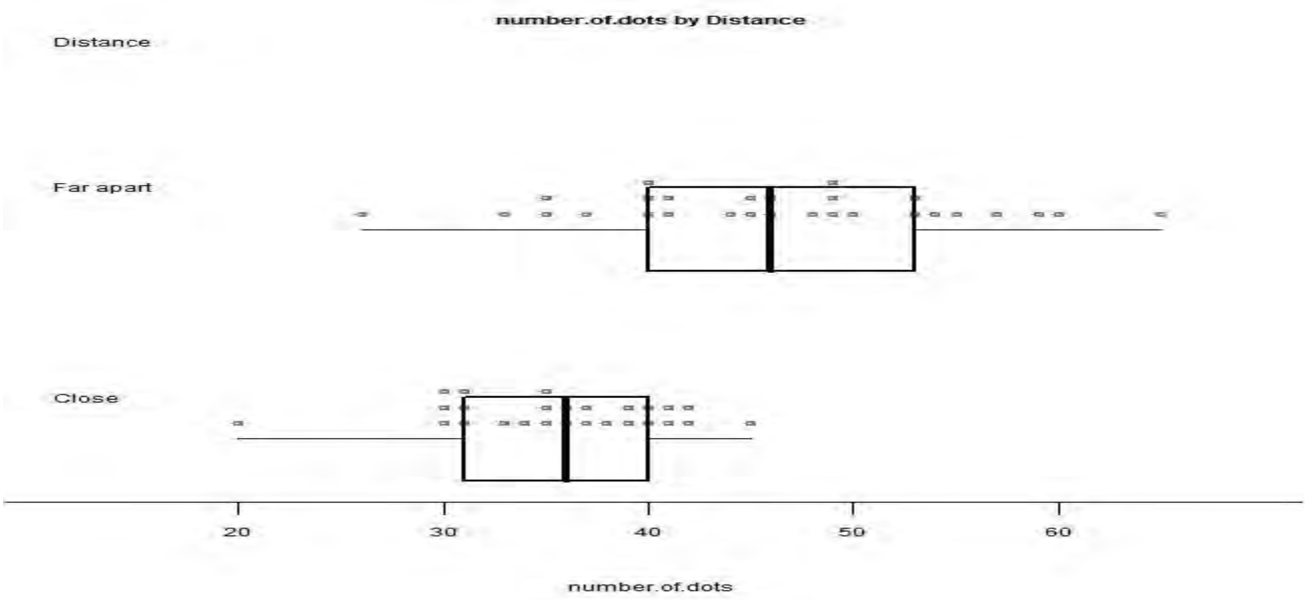
For the other group, the first slide I had was blank, so that it would give the students time to get ready. On the second slide I had 20 dots quite spread out. This slide was 3 seconds long. The third slide I wrote on it that they only had 5 seconds to write down their estimates. So the slides and how I ran the experiment was exactly the same for the both groups except for how the dots were laid out on the page.

I made the dots only visible for 3 seconds so that people couldn't count them, and only gave them 5 seconds to write down how many dots there were so they had to go with their first/quick guess.

I used a two Year 9 classes to carry out our experiment on. I explained to them what I was doing, like I was trying to see if they could guess how many dots were shown on the screen. I handed out pieces of paper to each student in the class with either A or B on them (I did this randomly so the students were randomly allocated to one of the groups). I didn't want them to write their names on their papers because I wanted it to be anonymous. My response variable was the student's estimate of the number of dots. My treatment variable will be having the same number of dots on two slides. On one slide I will have the dots bunched together and on another slide I will have the dots more spread out.

When I ran the experiment the first time, nearly all of the students in both groups guessed there were 20 dots. I realised that this was not a good number of dots to use - it was not too many and too easy to count roughly how many were shown and just round to a tidy number. So I did the experiment again, but used 43 dots on the slides (still with one slide/group with the dots close together, and one slide/group with the dots far apart).

The graphs and statistics for our data from the experiment is shown below:



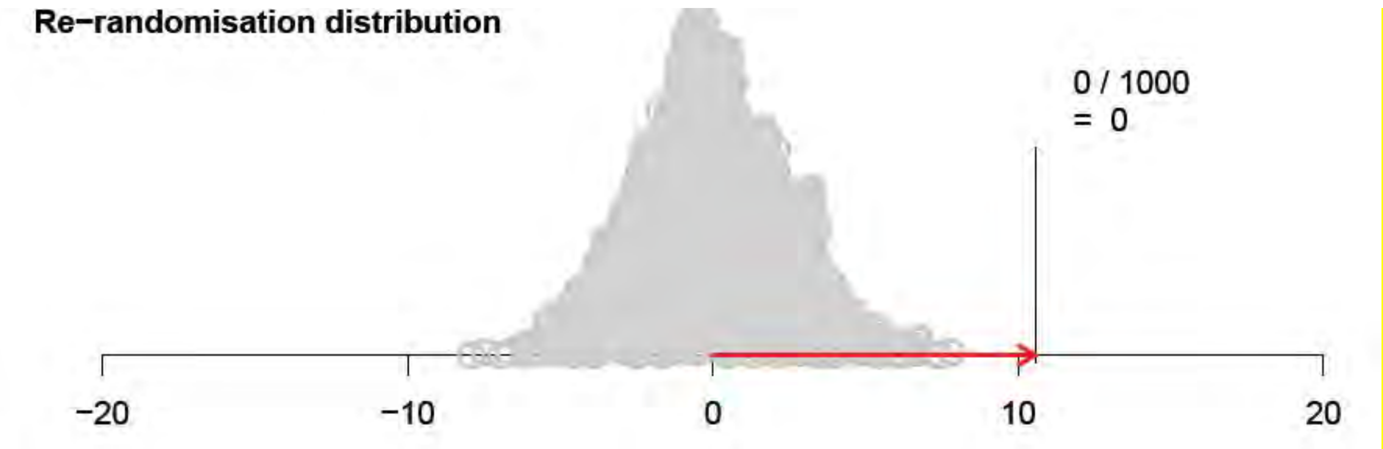
Summary of number.of.dots by Distance

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	Sample.Size
Close	20	31.5	36	35.69	39.75	45	5.335	26
Far apart	26	40.0	46	46.25	53.00	65	9.009	28

It's interesting that the estimates for the number of the dots from the far apart group are more spread out than the estimates for the dots from the close group. The box is wider and the standard deviation for the far apart group estimates of dots is higher – this shows me that people were less sure of the number of dots when they were displayed spread out than when they were displayed close together.

The difference of the medians of the two groups is 10 dots, whereas the difference of the means of the two groups is 10.56 dots. The distribution of the estimates for the far apart group is symmetric in shape, but the close group is slightly skewed. When I compare the features of the graphs for the two groups, they do not look like what I would expect if chance was acting alone.

I used the difference between the means of the two groups to look for evidence to answer my investigative question. I need to find out if it is likely to get a difference as big or bigger than 9.56 by chance alone. I used the randomisation test 1000 times to produce just by chance 1000 differences between the group means. The graphs and results produced from this method are shown below:



My observed difference of 9.56 only came up 0 times out of 1000. (2)

Because the probability is zero, this gives me really strong evidence that having the dots displayed far apart would cause estimates for the number of dots that tend to be higher than when they are displayed really close. It would be very unlikely that a difference as large as 9.56 could happen by chance alone, so something else must be working with chance to explain the effect – the difference of 9.56 dots between the groups. Because my experiment (when I re-designed it) went well and I minimised the effect of other variables on the estimates made, I can claim that having the objects (dots) spread out was what affected the number of objects (dots) estimated and caused estimates that tended to be higher. (3)

Appendix

Results:

DISTANCE	NUMBER OF DOTS
Close	36
Close	30
Close	35
Close	33
Close	20
Close	30
Close	37
Close	31
Close	41
Close	40
Close	45
Close	30
Close	34
Close	40
Close	35
Close	39
Close	31
Close	42
Close	38
Close	37
Close	31
Close	42
Close	39
Close	41
Close	35
Close	36

DISTANCE	NUMBER OF DOTS
Far apart	35
Far apart	40
Far apart	40
Far apart	33
Far apart	35
Far apart	26
Far apart	50
Far apart	37
Far apart	65
Far apart	45
Far apart	45
Far apart	49
Far apart	40
Far apart	60
Far apart	54
Far apart	59
Far apart	46
Far apart	57
Far apart	41
Far apart	53
Far apart	55
Far apart	48
Far apart	49
Far apart	44
Far apart	49
Far apart	53
Far apart	41
Far apart	46

	Grade Boundary: High Achieved
4.	For Achieved the student is required to conduct an experiment to investigate a situation using experimental design principles. This involves showing evidence of using each component of the investigation process. There is evidence of: • Posing an investigative question (1) • Defining and explaining the response variable (2) • Features of the data relevant to the experiment have been discussed (3) • A causal inference has been made (4). To reach Merit the student would need to describe in more detail the experimental plan (in particular the treatment variable).

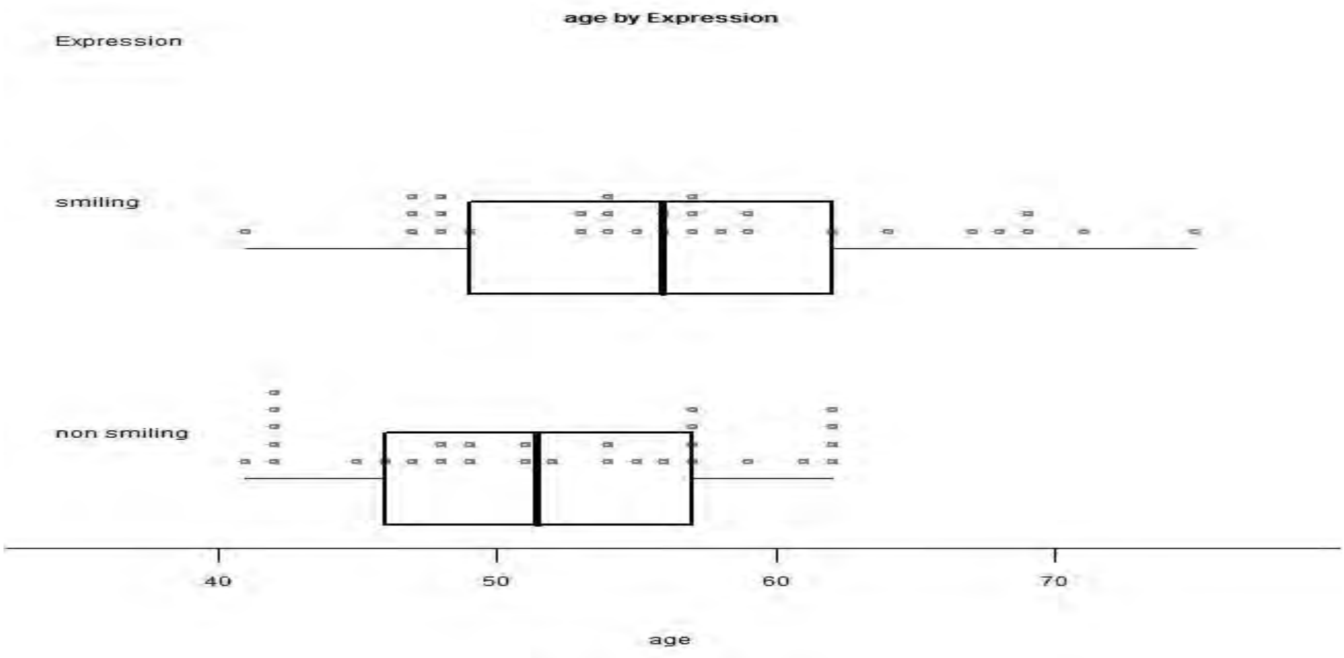
Student 4: High Achieved

My investigation was about whether appearance in particular the attractiveness of a person will affect the estimation of a person's age. We will be using two pictures; one picture with the person smiling and the other with the person not smiling. Because smiling makes you look more attractive, and attractiveness is linked to looking younger, I think we'll find lower estimates of the age from the photo of the person's facial expression (smiling). My experiment will be about whether slight changes (smiling or not smiling) in appearance can affect the estimation of a person's age? (1)

I had two pictures of the same person, George Clooney, at the age of 50. One had him smiling and the other had him not smiling but in both pictures his age was the same, this was to see if we could manipulate the outcome due to appearance. I had to make sure that the photos were as much the same as possible and the only difference was whether he was smiling or not. To ensure accuracy, I made sure the pictures were both taken at around the same period making sure he was the same age in both pictures. This was because if he was wearing different clothes or had his hair done differently, this could also affect the age people would estimate. (2)

I then required 2 classes to answer sensibly (silly answers were disregarded as well); the classes were two Year 9 maths class running at the same time of the day. Both classes were then combined and the students were then randomly allocated to two groups of 30. I did this by getting each student in the class to indicate by alternating between a 1 and a 2. I then asked all of the 1's to sit in room 24 and the 2's to sit in room 28. This meant I could compare two independent groups for the experiment. We shuffled up the pictures and then handed out the picture with George Clooney not smiling to one half and the picture with George Clooney smiling to the other half, the paper containing the photo also had instructions explaining to "Estimate this man's age in years" just in case the students weren't listening to the instructions we gave – Group 1 was given the photo smiling and group 2 was given the photo of not smiling which the students wrote down their estimates. We told the class not to look at each other's papers or talk.

After I had gathered the information I recorded it in a table on the computer.

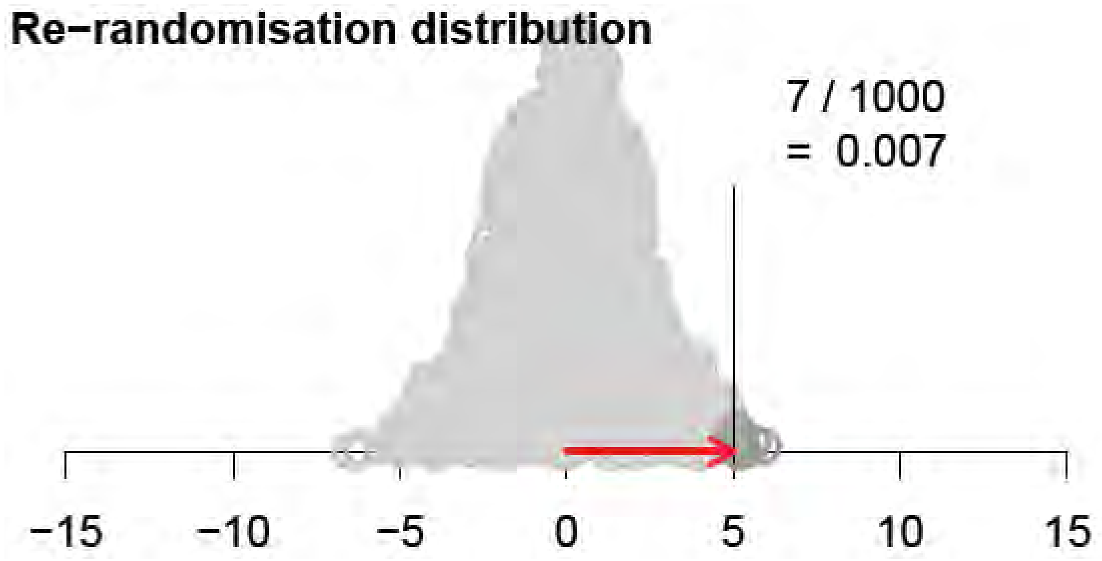


Summary of age by Expression

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	Sample.Size
non smiling	41.00	46.25	51.50	51.73	57.00	62.00	7.056	30
smiling	41.00	50.00	56.00	56.73	61.26	75.00	8.383	30

The box of smiling group is strenched out further than the non smiling group. There was only a difference of 1.3 years between the medians of the two groups, and no difference between the means of the two groups. In fact, the data looks like what I would expect if chance was acting alone - there is nothing I can see in the data that suggests the estimates for one group tend to be higher than the other, which was the point of my experiment. The results to me suggest people were making random guesses rather than being secure in their answers in both groups, but it is hard to tell with only 30 people in each group. (3)

I used the randomisation test on the means (the mean difference was 0 years) and the results are shown below:



As only 0.7% of the estimates produced by random allocation alone are at least as far from zero as the observed estimate, then the data provides strong evidence of a link between the two variables. This means that because the probability is low, it would be unlikely that a difference of 5 years could happen by chance alone, so something else must be working with chance to explain the effect. This might happen because the photos that we used in the experiment might be too similar and did not make much difference to the student's estimations.

I do not have enough evidence to say that the appearance of a person, specifically if they are smiling or not, can change the estimations of their age making them seem younger or older. This is due to two groups having a small difference between the two medians and a small difference between the two means. The randomisation test results show that the estimations of the ages 0.7% of the time are due to chance.
(4)

Some other variables that could have been changed for our experiment to obtain better data, we could of used a colour photo instead of a black and white photo, we could of used someone who looked younger to reduce chance of ridiculous results. Or perhaps people do find that people smiling look older than when they are not smiling, this could be brought on by many things e.g. while smiling wrinkles appear stronger than when not smiling.

Appendix

Results:

Expression	Age
Smiling	69
Smiling	75
Smiling	68
Smiling	47
Smiling	48
Smiling	47
Smiling	47
Smiling	56
Smiling	56

Expression	Age
Smiling	59
Smiling	64
Smiling	48
Smiling	49
Smiling	57
Smiling	54
Smiling	53
Smiling	57
Smiling	54

Expression	Age
Non Smiling	59
Non Smiling	57
Non Smiling	54
Non Smiling	61
Non Smiling	42
Non Smiling	48
Non Smiling	51
Non Smiling	48
Non Smiling	47

Smiling	54
Smiling	48
Smiling	53
Smiling	62
Smiling	55
Smiling	59
Smiling	67
Smiling	41
Smiling	69
Smiling	71
Smiling	57

Smiling	58
Non Smiling	62
Non Smiling	42
Non Smiling	46
Non Smiling	57
Non Smiling	54
Non Smiling	62
Non Smiling	62
Non Smiling	56
Non Smiling	55
Non Smiling	45

Non Smiling	42
Non Smiling	51
Non Smiling	62
Non Smiling	57
Non Smiling	57
Non Smiling	52
Non Smiling	49
Non Smiling	41
Non Smiling	42
Non Smiling	42
Non Smiling	49

	Grade Boundary: Low Achieved
5.	For Achieved the student is required to conduct an experiment to investigate a situation using experimental design principles. This involves showing evidence of using each component of the investigation process. There is evidence of conducting an experiment to investigate a situation using experimental design principles involves showing evidence of using each component of the investigation process (1). For a more secure Achieved grade the student would need to indicate how the surveys were distributed, provide detailed steps for the experimental plan and provide a more detailed answer to the question.

I decided to investigate how you could affect people's estimate of the length of time. The problem investigated was a scream and how this affected how long the student thought it took?

I used two classes of Year 9 students for the experiment. I am testing to see if a scream affects your ability to estimate the time. The students from the two different classes will be split into two groups. The experiment will be to play a scream and let the student listen to it. The specified time for both groups is 8 seconds. I will then ask the students to write down how long they thought the time was. To begin this experiment I will tell the class I will play a scream, then after I will ask them questions: before the experiment the students are not notified about what the experiment will be about. The response variable will be all the students' answers given in seconds.

The controlled variables are: one group will be timed for 8 seconds in silence while the other group will listen to the scream, I will ask them to keep quiet while the scream is being played. Both groups will be asked to write down how long their estimate of the time was. Variables that cannot be controlled are the student's ability to pay attention fully, the student's full participation, these points are made because some students do not give full attention in class, the noises caused by things outside of the classroom. The treatment variable is the scream.

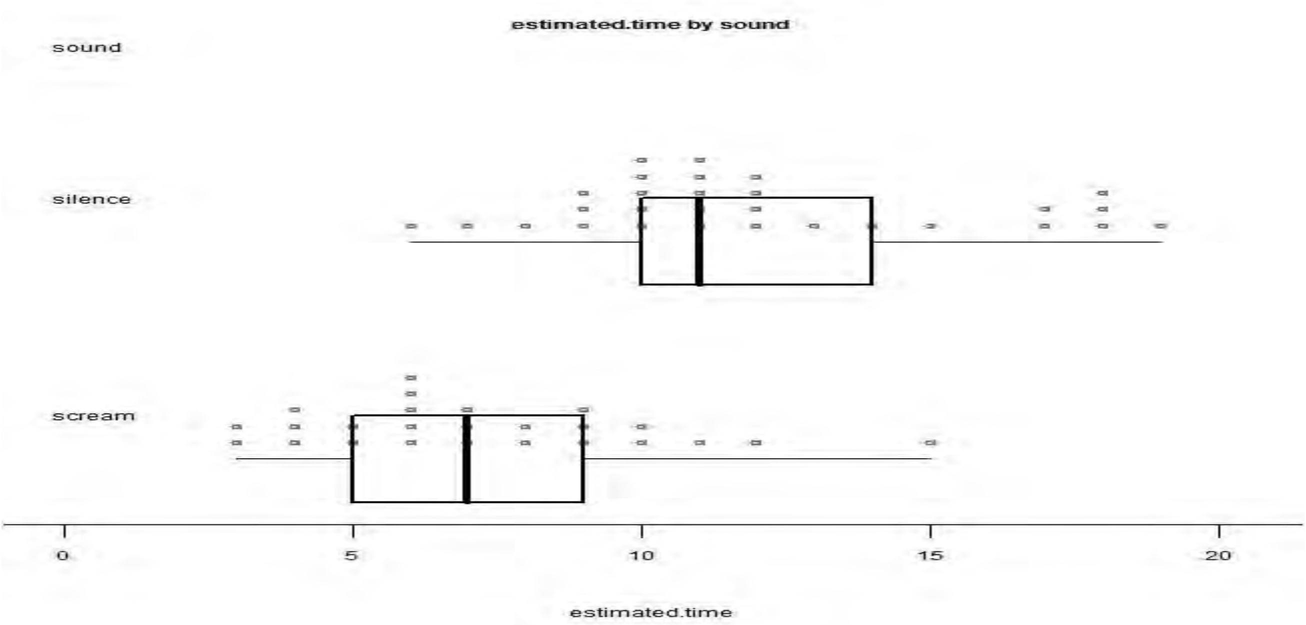
The two different groups will listen to different sounds.

1) Find one scream that lasts 8 seconds long.

2) Play the scream to the first group.

3) Time the second group for 8 seconds in silence.

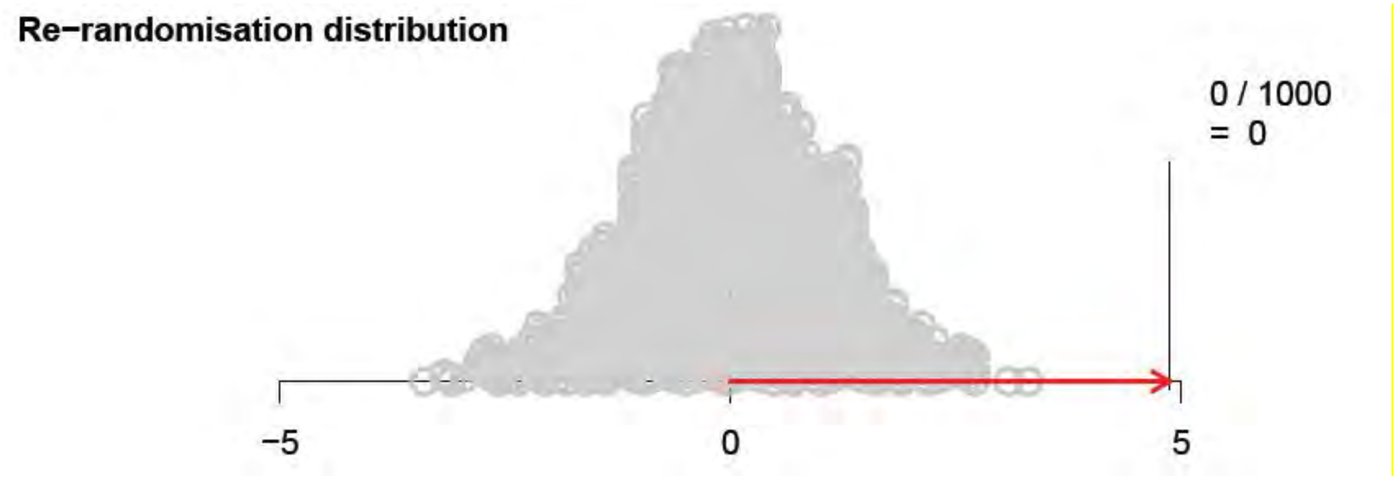
4) Retrieve and record the data (the data will be recorded by entering each individual student's data into the computer) (1)



Summary of estimated time by pitch

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	Sample.Size
Scream	3	5	7	7.20	9	15	2.944	25
Silence	6	10	11	12.07	14	19	3.535	29

The data from the groups do look like what I would expect to see if chance was not acting alone in the experiment. The estimates for the lengths of the scream are a lot lower than those from the timed silenced group with a difference between the two means of the groups of 4.87 seconds. I used the randomisation tool on iNZight with the means of the two groups.



As the estimates produced by random allocation of 0% are at least as far from zero as the observed estimate, then the data provides very strong evidence of a link between the two variables. A difference of 4.87 seconds between the two groups came up zero times out of 1000 re-randomisations. The results show that it would not be possible to find a difference of 4.87 seconds for the means of the two groups (scream and silence) could occur by chance alone. However, something else could also be happening. So I don't think that the scream that affects the estimates for the length of the time. A factor that could have affected the data would be the clock in the room, although it was taken down, this could have given a good indication to the class the experiment had to do with time. Secondly students are more likely to overestimate time when they are silent than when they hear noise. (1)

Appendix

Results:

SOUND	ESTIMATED TIME
Scream	6
Scream	12
Scream	3
Scream	5
Scream	5
Scream	7
Scream	9
Scream	3
Scream	4
Scream	6
Scream	9
Scream	10
Scream	11
Scream	15
Scream	4
Scream	6
Scream	8
Scream	4
Scream	7
Scream	9
Scream	6
Scream	8
Scream	7
Scream	6
Scream	10

SOUND	ESTIMATED TIME
Silence	10
Silence	11
Silence	17
Silence	18
Silence	11
Silence	7
Silence	12
Silence	11
Silence	18
Silence	17
Silence	19
Silence	9
Silence	10
Silence	13
Silence	14
Silence	12
Silence	11
Silence	10
Silence	9
Silence	10
Silence	11
Silence	10
Silence	9
Silence	6
Silence	8
Silence	12
Silence	18
Silence	15

			Silence	12

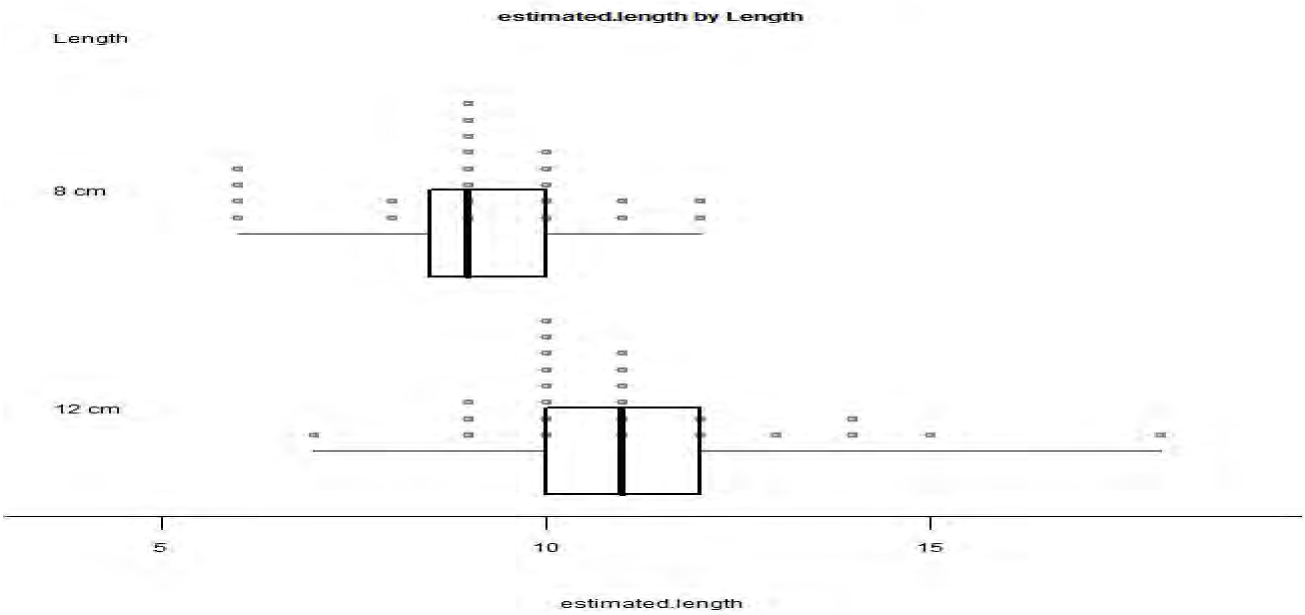
	Grade Boundary: High Not Achieved
6.	To achieve the standard the student is required to conduct an experiment to investigate a situation using experimental design principles. This involves showing evidence of using each component of the investigation process. The student has posed a question, attempted to carry out an experiment and answered the question (1). To be awarded Achieved there needs to be more detail in the experiment plan, indications of how the treatment was applied to the groups and further discussion of the displays and measures.

Student 6: High Not Achieved

I chose to investigate what might affect people's estimation of length. I wonder if students can correctly estimate the length of two lines which are a different length?

I carried out this experiment by preparing two lines on two separate pieces of paper. One paper had an eight centimetre line and the other had a twelve centimetre line. Both lines were drawn horizontally. The students who did our experiment were two year nine maths classes. The classes were chosen for us so we could get some silly answers.

The overall design of my experiment will involve a comparison of two independent groups. Students will be given only one of the two lines. The response variable that I have chosen is the estimate of the length of the line. The only difference between the two lines is that one line is eight centimetres long and the other had a twelve centimetres long both were drawn horizontally. The student will estimate the length of one of the lines (depending on which group they are in), using no ruler and not communicating with others. The students will be asked to estimate the length of the line to the nearest centimetre.

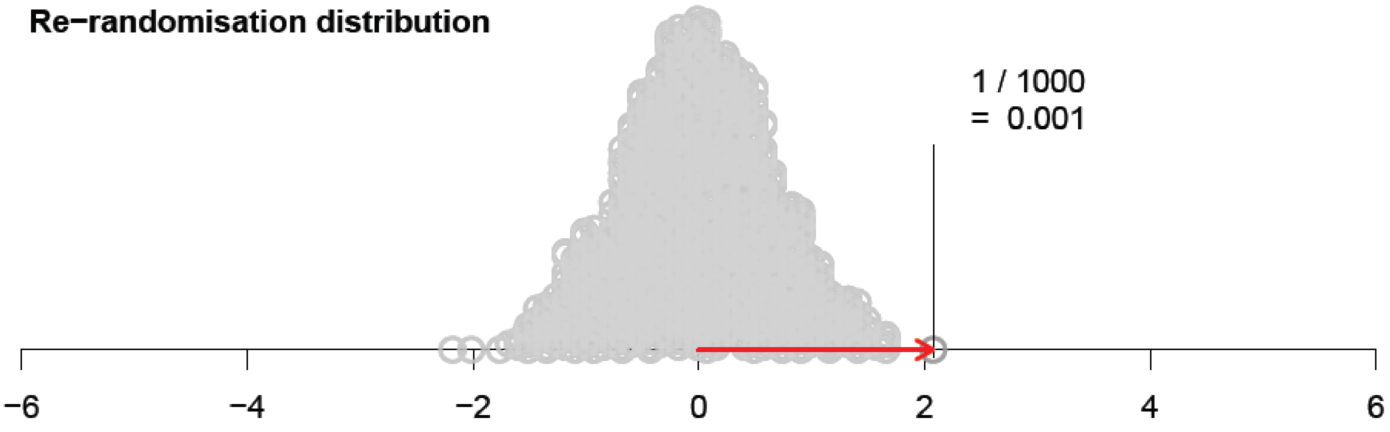


Summary of length by estimated length

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	Sample.Size
12 cm	7	10	11	11.120	12	18	2.279	25
8 cm	6	8.5	9	9.043	10	12	1.770	23

From my data I can see that there is a difference between the medians and means of each of the groups. The students from the two groups mostly recognised how long the line wa. The distributions of the estimates for the lengths in each group are very similar, with most estimates either 10 or 11cm. The spread of the estimates of the lengths for each group are pretty much the same (the boxes are the same width, and the standard deviations are similar).

I did the randomisation test using the means. This will take the values from the groups and randomly re-assign them to one of the two groups and calculate the difference between the re-randomised group medians 1000 times. The results are below:



A difference of 2.077 cm came up once out of 1000 for the re-randomised differences.

I can therefore conclude that students can correctly estimate the length of two lines which are a different length (1)

My experiment was not designed well. The students could clearly and confidently identify the two lines as being approximately 10cm. You can see that a lot of the students either got the answer or very close to it. If there is anything about my experiment that I did that I could have improved on, it would be to make the lines a more difficult measurement. I think 8 cm and 12 cm were too obvious – it would have been better to make the lines longer so the students were not sure about how long the lines were.

Appendix

Questionnaires/tools used:

Sighted by teacher but not provided.

Results:

Treatment	Estimate of length (cm)
8cm	10
8cm	10
8cm	9
8cm	10
8cm	9
8cm	9
8cm	8
8cm	6
8cm	10
8cm	12
8cm	9
8cm	10
8cm	11
8cm	6
8cm	9
8cm	6
8cm	11
8cm	9
8cm	12
8cm	9
8cm	6
8cm	8
8cm	9

Treatment	Estimate of length (cm)
12 cm	10
12 cm	10
12 cm	10
12 cm	11
12 cm	15
12 cm	18
12 cm	11
12 cm	7
12 cm	12
12 cm	11
12 cm	10
12 cm	14
12 cm	11
12 cm	9
12 cm	10
12 cm	13
12 cm	14
12 cm	12
12 cm	11
12 cm	10
12 cm	9
12 cm	10
12 cm	11
12 cm	10

12 cm	9

<u>Experiment notes:</u>

Even though students were told to estimate to the nearest mm, all the students wrote down whole number estimates.