Tricky questions – high excellence exemplar

<table>
<tr><th colspan="3">Introduction (Problem and purpose)</th></tr>
<tr>
<td>Achieved
Posing an investigative question about a given experimental situation

e.g. The student has clearly stated what they are going to investigate (the experimental situation)

e.g. The student has posed a causal relationship question that can be investigated by conducting an experiment</td>
<td>I investigated possible bias in questionnaires from using anchors. I wanted to find out if you could get people to give higher answers for a question by using an anchoring question before it. This would be important to take into account when using questionnaires to collect data so you don't unknowingly influence answers, or something that might be used in questionnaires to trick people into giving a certain answer.

I spent some time researching each of the different types of questionnaire bias given in resource A, and it seemed to me that anchoring effect would be interesting to investigate because of the applications of this to everyday life. In my research, I found examples of how people use anchoring bias when selling cars to persuade people into paying more. For example, the sale person might tell a person that the car is worth $30 000 and then offer it to them for $15 000. If the person buying the car doesn't really know how much the car is worth, they will use the $30 000 amount as an anchor to decide if they are getting a good deal, so they might think they were getting a 50% discount.

The research about anchoring bias shows that when people are uncertain about something, they use whatever information is available to help them decide, even if the information is not valid or reliable. In an experiment by Amos Tversky and Daniel Kahneman (see http://en.wikipedia.org/wiki/Anchoring), they showed that when asked to guess the percentage of African nations that are members of the United Nations, people who were first asked "Was it more or less than 10%?" guessed lower values (25% on average) than those who had been asked if it was more or less than 65% (45% on average).

I wanted to do a similar experiment, so my investigative question was "will higher estimates be given for the number of people who live in Venezuela if a larger number is used for the anchor?" I used Venezuela for my experiment as I read about someone else using this as a different context (see http://youarenotsosmart.com/2010/07/27/anchoring-effect/) and I thought that it was a country not many people would know exactly how many people lived there, and for the experiment to work people have to be unsure about the answer or amount for the anchor to have an effect. The research suggested that the higher the number I used in the anchor, the higher the estimates of the number of people who live in Venezuela would be, and so this is what I expected to find in my experiment.</td>
<td>Merit
Linking to the context
Explaining relevant considerations
Supporting findings with evidence

e.g. The student has clearly stated what they are going to investigate (the experimental situation)

e.g. The student has posed a causal relationship question that can be investigated by conducting an experiment

e.g. The student has presented a hypothesis or prediction for their experiment, which is justified using research findings or a well-reasoned explanation

Excellence
Integrating statistical and contextual knowledge
Reflecting about the process
Discussing how possible sources of variation were dealt with
Considering other relevant variables

e.g. The student has demonstrated a depth of understanding of the background contextual knowledge to develop the investigative question
e.g. The student has presented a hypothesis or prediction for their experiment, which is justified using research findings or a well-reasoned explanation</td>
</tr>
</table>

Method (Plan)

Achieved

Planning the experiment using experimental design principles

e.g. The student has identified the type of experiment they will conduct

e.g. The student has described what the experimental units will be for the experiment

e.g. The student has identified the response variable for the experiment and how it will be measured

e.g. The student has identified the treatment variable, including the treatment levels and groups, and how this variable will be manipulated for the experiment

e.g. The student has identified or given details about any measurement tools designed for the experiment

e.g. The student has identified other sources of variation that could affect the results of the experiment

e.g. The student has described how the treatments will be allocated to the experimental units

e.g. The student has described how the data from the experiment will be collected

I used a comparison of two independent groups for my experiment, where one group was given a high anchor and one group was given a low anchor. The groups had to be independent because if I gave someone both anchors then they might realise what the experiment was about and I wanted their behaviour to be natural. I used two Maths classes for my experiment (57 students from Year 10 and Year 12 Maths classes), which was convenient because they were on at the same time as my statistics class. The experiment was conducted during my Statistics lesson period 4 on Wednesday. I used single blinding, where the participants didn't know which treatment they were getting. In fact, I concealed the fact that it was an experiment at all by presenting the questionnaire as a general knowledge survey.

The response variable was the estimate for the number of people in Venezuela (in millions). I decided to use millions for the unit as this generally what populations for different countries as measured in. The treatment variable was the number used for the anchor. I had two treatment groups: for one group the anchor was 60 million, for the other group the anchor was 6 million. I used these two numbers as the population of Venezuela is around 29 million. I used 6 million as the other number to see if I could get people to estimate a much lower population size (I wanted it to be quite far away from the real population size). I used the same idea with setting the other number at 60 million.

I created two different questionnaires for my experiment – these are provided in the appendix of this report. I included an introduction for the questionnaire used in the experiment that said it was a general knowledge survey, and asked a couple of other questions in the questionnaire (What is your gender? What year level are you in? How many continents are there in the world?) so that people would not guess the point of my experiment. In my questionnaire, I decided to ask people to estimate of the number of people in Venezuela to the nearest million because I was confident this would still give me enough variability in the estimates and it would be easier for people to answer the question. The two versions of the questionnaire are exactly the same, except for the number used for the anchor in the question before the one that asks people to estimate the number of people in Venezuela (e.g. Is the number of people in Venezuela bigger or smaller than 6 million?). I found in my research about questionnaire bias that the order of questions can influences responses, so it was important that I didn't change the order of any questions in the questionnaire used for the experiment.

For this experiment, it was important that people didn't realise there were two different versions of the questionnaire. Before we gave them to students to fill out, we used random numbers to randomise the order of the questionnaires into a pile to hand out. In this way, we would be randomly allocating students to one of the two treatment groups when we gave out the questionnaires to complete.

There are some factors I couldn't control for my experiment, like:

- whether people already know the population of Venezuela (maybe people who had been travelling or international students from South America)
- whether people would take the survey seriously and not give silly answers
- whether people were aware of the anchoring bias (which maybe students doing psychology might know about)
- whether some people had better general knowledge than others

Merit

Linking to the context
Explaining relevant considerations
Supporting findings with evidence

e.g. The student has identified the type of experiment they will conduct

e.g. The student has described what the experimental units will be for the experiment

e.g. The student has justified how the response and treatment variables (including levels and groups) were defined for the experiment

e.g. The student has identified or given details about any measurement tools designed for the experiment

e.g. The student has identified other sources of variation that could affect the results of the experiment

e.g. The student has described how the treatments will be allocated to the experimental units

e.g. The student has described how the data from the experiment will be collected

e.g. The student has described how the data from the experiment will be recorded

Excellence

Integrating statistical and contextual knowledge
Reflecting about the process
Discussing how possible sources of variation were dealt with
Considering other relevant variables

e.g. The student has identified the type of experiment they will conduct

e.g. The student has described what the experimental units will be for the experiment

e.g. The student has justified how the response and treatment variables (including levels and groups) were defined for the experiment

e.g. The student has identified or given details about any measurement tools designed for the experiment

e.g. The student has used contextual knowledge to

	By randomly assigning people to one of the two treatment groups, I attempted to balance the possible effect of these variables on the estimates across the two groups. For example, people who may be better at estimating the population of Venezuela should be balanced across both groups and should not significantly affect the results for one group only. The other variables I controlled for my experiment were: • giving the same instructions to people about completing the questionnaire • both groups doing the experiment at the same time of the day • same test conditions used for completing the questionnaire • an independent person carrying out the experiment (the teacher of the class) so that I didn't influence the results if I was the handing out the questionnaires I didn't tell the teacher that the survey was an experiment – I told them it was just a survey to collect data for analysis for our Statistics class. In this way, they were not aware of the two treatment groups. The experiment was done at the beginning of the lesson by the teacher reading out the instructions (which are in the appendix). I entered the responses from each questionnaire into a spread sheet and made notes about any responses that were not clear (see appendix).	***develop a list of relevant variables that could affect the response variable and has used statistical knowledge to develop a plan for their experiment that clearly describes how these sources of variation will be controlled or balanced*** ***e.g. The student has described how the data from the experiment will be collected*** ***e.g. The student has described how the data from the experiment will be recorded***

Results (Data and analysis)

Achieved

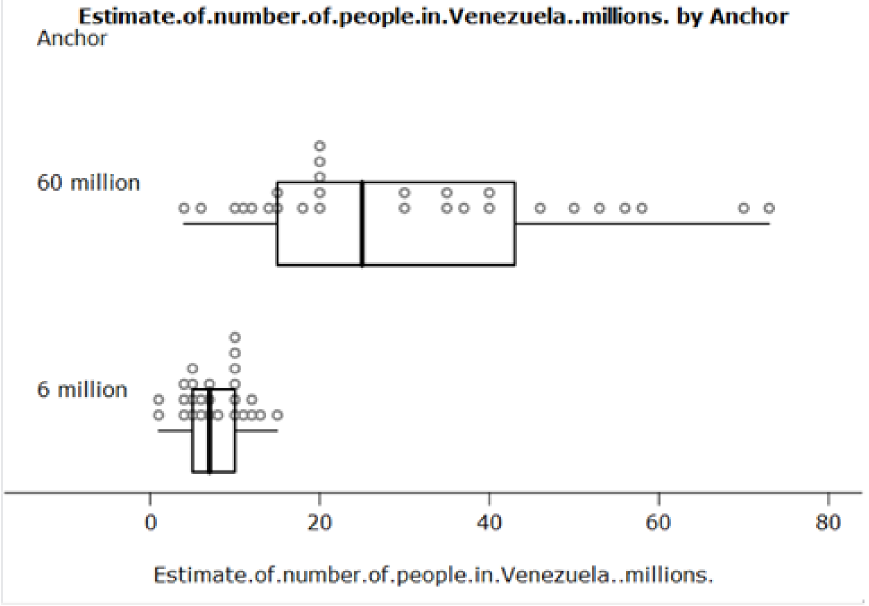
Selecting and using appropriate displays and summary statistics

e.g. The student has produced the displays and statistics appropriate to the design of their experiment

e.g. The student has described key features of the displays and statistics relevant to the experiment

e.g. The student has selected an appropriate statistical method to obtain evidence to answer their investigative question

The raw data collected from the experiment is included in the appendix of this report.
To analyse the data I used iNZight to generate the following dot and box-and-whisker plots and summary statistics.



Summary of Estimate.of.number.of.people.in.Venezuela..millions. by Anchor

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	Sample.Size
6 million	1	5	7	7.68	10.0	15	3.682	25
60 million	4	15	25	30.64	41.5	73	19.311	28

The first thing I noticed about the displays when comparing the two groups is that the estimates for the number of people in Venezuela do not look like what I would expect if chance was acting alone. For the amount of data I have, under chance alone, I would expect the data for both groups to look similar – the boxes would be sitting roughly in the same place, they would be roughly the same size, and the shape of the distribution of the data would be similar.

However, there is a difference of 18 million between the two treatment group medians (the treatment group which used an anchor of 60 million had a median estimate of 25 million whereas the treatment group with the 6 million anchor had a median estimate of 7 million). This difference, especially with the non-overlap of the boxes in the box-and-whisker plot, suggests that the use of an anchor of 60 million would result in estimates which tend to be higher than when a 6 million anchor is used.

Merit

Linking to the context
Explaining relevant considerations
Supporting findings with evidence

e.g. The student has produced the displays and statistics appropriate to the design of their experiment

e.g. The student has described key features of the displays and statistics relevant to the experiment and has linked features of the data to features of the experiment

e.g. The student has selected an appropriate statistical method to obtain evidence to answer their investigative question

e.g. The student has justified their choice of statistical method and explained how the method allows them to detect the existence of the causal relationship

Excellence

Integrating statistical and contextual knowledge
Reflecting about the process
Discussing how possible sources of variation were dealt with
Considering other relevant variables

e.g. The student has produced the displays and statistics appropriate to the design of their experiment

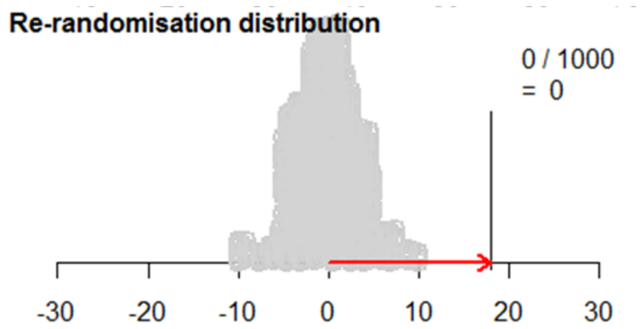
e.g. The student has explored the data collected from the experiment and used relevant features of the data to consider further factors or effects

e.g. The student has selected an appropriate statistical method to obtain evidence to answer their investigative question

e.g. The student has justified their choice of statistical method and explained how the method

Also, the variability in these estimates for the 60 million anchor group was much smaller than the variability for the 6 million anchor group – you can see this in the box-and-whisker plot, as the box for the 60 million anchor group is much wider. This shows that people who were in the 60 million anchor group were more inconsistent in their estimates, and most gave values quite a bit lower than the anchor of 60 million. People who were in the 6 million anchor group gave estimates pretty close to the anchor of 6 million and their estimates are roughly unimodal and roughly symmetrically distributed around the median of 7 million. (I suspect that the single digit nature of the 6 million anchor, perhaps together with the zero lower bound, somehow signalled that an estimate should be given to within a single digit of the anchor whereas this was not the case for the two digit 60 million anchor.) The distribution of the estimates for the 60 million anchor group is right skewed. The mean and median for the estimates of the 6 million anchor group are very similar, whereas there is a 5 million difference between the mean and median for the estimates of the 60 million anchor group. The median and mean values for the 60 million group are quite different because the estimates from this group are skewed.

I have decided to focus on the difference between the medians of the two groups (I could have also used the means) 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 18 million by chance alone, I mimicked the data production process of random assignment to two groups but with only chance acting. I did this using a computer package. Each observed value was randomly re-assigned (re-randomised) to one of the two groups and the difference between the re-randomised group medians 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:



Not one of the 1000 re-randomised differences was equal to or higher than the observed difference between the medians of 18 million. A difference as big as 18 million was highly unlikely to have happened by chance alone.

allows them to detect the existence of the causal relationship

Discussion (Conclusion and reflection)		
Achieved Making appropriate formal statistical inferences ***e.g. The student has used the statistical method selected to make a correct inference about the causal relationship investigated*** Communicating findings in a conclusion. ***e.g. The student has clearly communicated each component of the investigation process***	Using the result from the randomisation test, I have incredibly strong evidence that the use of an anchor of 60 million would cause estimates that tend to be higher than when a 6 million anchor is used. This is because when I compared the observed difference between the group medians (18 million) to the distribution of re-randomised differences, a difference of 18 million or higher never came up in a 1000 re-randomisations. This shows that in this experiment it would be incredibly unlikely that a difference as large as 18 million could happen by chance alone. It is this test result that provides me with the incredibly strong evidence that chance was not acting alone in this experiment but something else, namely the anchor effect, was acting along with chance to create the observed difference of 18 million. I thought that the experiment turned out reasonably as planned, although not everyone completed the survey correctly, so I couldn't use all the results (see my notes in the appendix). However, there were not a large number of incomplete or invalid responses, so this shouldn't have affected my data too much. I could have also the teacher to check the questionnaires as they were handed back so that responses could be clarified. The results from the randomisation test, and the fact that my experiment was well- designed and executed, means that I can claim that an anchor of 60 million would cause estimates that tend to be higher than when a 6 million anchor is used. The results from my experiment are also consistent with other research studies for anchoring bias, and confirm that the anchoring bias could be present in a variety of situations. My results are also important in terms of how they apply to questionnaire design, and the importance of making sure that questions are not used that may influence people's answers to other questions in the questionnaire. I also assumed that the two year levels (Year 10 and Year 12) would be similar in their ability to be able to estimate the number of people in Venezuela. However, when I was entering the data from the questionnaires, I noticed that the Year 10 estimates seemed to be higher, regardless of which treatment group they were in. I think that Year 10 students are more likely to be silly and just guess higher numbers and also that being younger perhaps they are more likely not to know populations of countries. I decided to see how many Year 10 and 12 students ended up in each of the treatment groups when I randomly allocated them, and found 60% of the Year 10 students were in the 60 million treatment group. If I was to conduct this experiment again, I could use blocking as part of my experimental design. I could block on the year level, and this would allow me to test out if there was difference between the estimates given by Year 10 and Year 12 students. This would involve randomly assigning the 60 million and 6 million anchors within each of the year levels, to make sure each year level had a balance of 60 and 6 million questions. I wonder how much the numbers I used for each of the treatment groups (the numbers for the anchors) affected the estimates. When I looked at the data for the two treatment groups, it seemed that people were not confident estimating the number of people in Venezuela to be as high as 60 million (the median estimate was 23 million). Maybe if I had used a value like 100 million	**Merit** Linking to the context Explaining relevant considerations Supporting findings with evidence ***e.g. The student has assessed and interpreted the strength of evidence for the inference about the causal relationship investigated*** ***e.g.The student has linked the design and conduct of their experiment with their results when communicating their overall findings*** **Excellence** Integrating statistical and contextual knowledge Reflecting about the process Discussing how possible sources of variation were dealt with Considering other relevant variables ***e.g. The student has discussed how their findings relate to the other research findings presented in the introduction*** ***e.g. The student has reflected on how the experiment was conducted, identified key issues in their design, and explained how these design issues affect their findings*** ***e.g. The student has considered broadening the experimental situation in their discussion of their findings***

	they would have ignored it because it would have been unrealistic. Would I have got the same result (my conclusion that a larger number for the anchor would result in (cause) estimates that tend to be higher) if I used treatment values of 30 million and 20 million? So perhaps anchors can influence people's responses to answers, but only if they are a certain value. It would be interesting to repeat the experiment to see how far apart the numbers I use for the two treatment groups have to be to see an effect in the estimates for the number of people in Venezuela. It would also be interesting to see what people's estimates would be for the number of people in Venezuela without any anchor. The distribution of the estimates for the number of people in Venezuela (both the shape and spread) also suggests that the use of an anchor of 60 would result in (cause) estimates that tend to be more varied than when a 6 million anchor is used. I could investigate by using statistical tests to see whether there is evidence of a difference of spread for the estimations due to the intervention of the anchors.	

Appendix (Raw data and experiment notes)	
Achieved Conducting the experiment *e.g. The student has collected data from the experiment* *e.g. The student has made notes about the conducting of their experiment* *e.g.The student has included any tool (i.e. questionnaires or tasks) or information or instructions used to conduct the experiment*	Statement read out to the class: I am going to give you a general knowledge survey to complete. Please answer the questions as best you can and place your survey face down when you have finished for me to collect. Keep your eyes on your own survey and do not talk – just like a test. Questionnaires used: **General knowledge survey A** What year level are you? ________ What is your gender? (please circle answer) male female How many continents are there in the world? _________ Is the number of people in Venezuela bigger or smaller than 6 million? (please circle answer) bigger smaller Estimate, to the nearest million, the number of people in Venezuela ______________ **General knowledge survey B** What year level are you? ________ What is your gender? (please circle answer) male female How many continents are there in the world? _________ Is the number of people in Venezuela bigger or smaller than 60 million? (please circle answer) bigger smaller Estimate, to the nearest million, the number of people in Venezuela ______________ Raw results:

	Student	Year level	Gender	Number continents answer	Anchor	Estimate of number of people in Venezuela (millions)	Data entry comments
	1	12	m	5	6 million	6	
	2	12	M	5	6 million	4	
	3	12	F	7	6 million	5	
	4	10	F	7	60 million	11	
	5	10	f	7	60 million	35	
	6	12	M	6	6 million	8	
	7	12	M	7	60 million	30	
	8	10	F		60 million		No answers given
	9	10	M	8	60 million	73	
	10	12	M	7	60 million	35	
	11	10	M	5	60 million	40	
	12	10	F	6	6 million	7	
	13	10	F	7	60 million	20	
	14	12	F	7	60 million		No estimates given
	15	12	M	7	6 million	10	
	16	12	M	6	6 million	10	
	17	10	M	5	6	6	Gave 5.9 as answer

					million		
	18	10	M	6	60 million	20	Written as 20,000,000
	19	12	M	7	6 million	10	
	20	10	F		6 million		No estimates given
	21	10	M	7	60 million	56	
	22	10	f	7	60 million	15	
	23	10	f	7	60 million	15	
	24	10	M	7	6 million	11	
	25	12	M	6	6 million	13	
	26	10	F	7	60 million	14	
	27	12	F	5	60 million	20	
	28	12	M	7	6 million	1	1000 written
	29	12	F	7	6 million	4	
	30	12	M	1000	6 million	10	10000000 given
	31	10	F	7	6 million	5	
	32	12	F	5	60 million	50	
	33	10	m	134	60 million	30	
	34	10	F	7	6 million	5	
	35	10	f	7	60 million	12	
	36	12	f	7	6	12	

				million		
37	10	M	5	60 million	70	
38	12	M	1	6 million	1	1000 given
39	12	F	6	60 million	6	
40	10	F	7	60 million	20	
41	10	F	7	60 million	18	
42	10	M	12	60 million	53	
43	10	F	7	6 million	7	
44	12	F	54	60 million	20	
45	12	M	6	6 million	12	
46	10	F	6	6 million	7	
47	10	F	7	6 million	4	
48	10	F	5	6 million	15	
49	12	M	7	60 million	46	
50	10	M	216	6 million	10	
51	12	M	5	60 million	40	
52	12	F	7	6 million	5	
53	12	F	5	60 million	58	
54	12	M	7	60 million	10	
55	10	F	7	60	4	Answer given as

				million		4,000,000
56	12	F	7	6 million	10	
57	12	M	5	60 million	37	Written as 13 37 mil

<u>Experiment notes:</u>

The classes were pretty quiet and it didn’t look like anyone copied from each other or talked about the questionnaire.
I ended up with a few questionnaires left over, with more of the “6 million” question.