STATISTICS & DATA SCIENCE EDUCATOR

Experiencing experimental design principles: The Elevens trick

Barbara Costelloe
Sacred Heart Girls' College, New Plymouth, New Zealand
Email: bco@shgcnp.school.nz

Overview of lesson

In this lesson students experience an experiment first-hand. Students then reflect on experimental design principles in evaluating the experiment. The experiment is conducted as a comparison of two independent groups.

Learning objectives

- Develop understanding of how an experiment is designed through participation
- Reflect on the validity of results and other explanations
- Critically appraise the experiment in terms of experimental design principles and issues that arose.

Suggested age range

The activity is recommended for Year 13 students.

Time required

Approximately one to two periods (hour sessions): one session to conduct the experiment and a further session for reflection on the experience is recommended.

Keywords

experiment, random allocation, treatment, control, response, experimental units

Introduction

The idea for the lesson came from a trick a family friend showed me many years ago. The trick was how to multiply large numbers by 11 in one step simply by adding adjacent digits. I realised that teaching the trick to half a class only and then giving them all a traditional times table test like they had in primary school would make an experiment that was quick and easy to run in a regular senior classroom.

Because the students themselves are the experimental units, the factors affecting their ability to perform the test should be self-evident. The experiment also gives a context to discuss maths anxiety and brain-based learning with students, which could fuel some creativity for the types of experiments they could conduct themselves.

I use this lesson as an introduction to Experiments in Year 13. The intention is that students experience being the experimental units and then have the opportunity to reflect on the effect of the treatment, the influence of other sources of variation and possible improvements to the experiment. It does not matter if the results of the experiment are not perfect as they will fuel critical analysis of the planning stage. Following this introduction to experiments, it is intended that students would design and conduct their own experiments.

Illustration: Timothy Costelloe

Lesson outline

The lesson is conducted over two sessions. In Session One the students experience the experiment as experimental units and analyse the outcomes of the experiment. In Session Two they reflect on their analysis and the factors that could have influenced the results. Finally, they are introduced to experimental design principles in cognisance of the experiment that they have just participated in.

Session One

1. The experiment

At the start of the experiment it is important not to tell the class what the experiment is about as the control group should have no prior warning of what is coming.

> One year I told the students they would all be getting a test on the 11 times table and one student in the control group taught everyone else the traditional multiplication method while I taught my 11s trick. Consequently, there was no evident treatment effect!

In order to randomly allocate the treatment to the experimental units, I count out playing cards so that each student gets one and ensure that there is an even mix of red and black cards. After shuffling the cards, I offer each student the deck of cards face down to select a card in order to randomly assign them to the red or the black group. Next, I ask who has the next birthday in the class. This student randomly chooses a card from the remaining pack of cards to determine which colour will be the control group and which will be the treatment group.

> I don't give any explanation about why I am using these two techniques at the time, but they are fully discussed in session two, as part of the experimental design principles, when we look at why we need to randomise and how best to do this.

I take the students in the treatment group outside the classroom and fairly quickly explain how to multiply a large number by 11 using the "11s trick": To get the answer simply write down the last digit, then add each pair of consecutive digits from right to left and lastly write the front digit at the front of the answer. The trick is a little more challenging when sums add to more than 10 as it requires "carrying" to the next digit (see Figure 1).

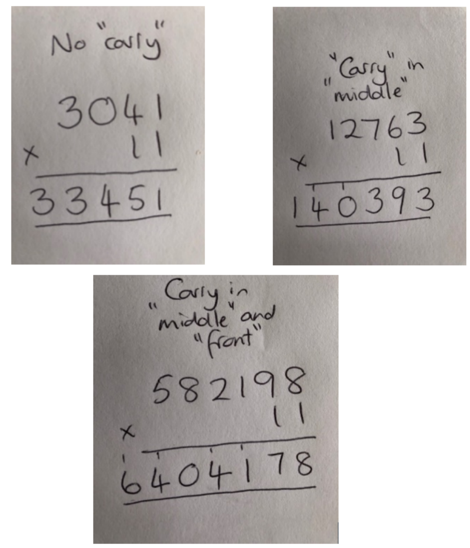
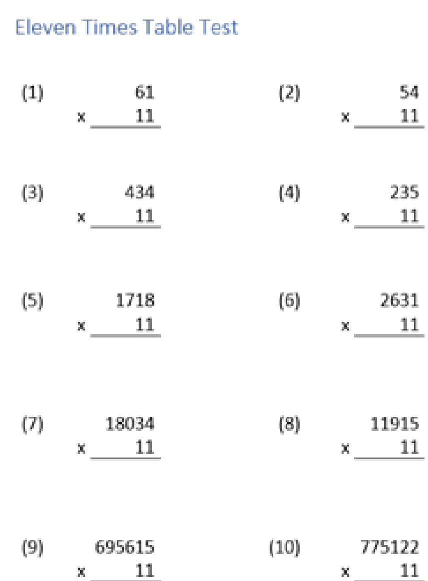


Figure 1: Examples of explanations of the "11s trick" given to treatment group by teacher

Students are handed out test papers (see Figure 2) face down and told that they have two minutes to complete the 10 questions without calculators. I time two minutes on my cell phone.


Eleven Times Table Test

(1) 61 x 11 (2) 54 x 11

(3) 434 x 11 (4) 235 x 11

(5) 1718 x 11 (6) 2631 x 11

(7) 18034 x 11 (8) 11915 x 11

(9) 695615 x 11 (10) 775122 x 11

Figure 2: Example of a test paper

2. Collecting the data

At the end of the two minutes, students are asked to swap papers and I write the answers on the board for them to mark the papers. I use this opportunity to share the 11s trick with everyone. I then open up a spreadsheet and display this on the data projector for students to check as I collect the results from the class. Students call out whether they were in the black or red group and their score for the number of correct calculations.

We discuss whether this is an efficient way to collect the data or whether it would be better to take in all the papers and enter the data into the spreadsheet individually.

3. Analysing the results

I use the software, *NZGrapher* (www.grapher.nz) to produce boxplots and summary statistics for the results (Figure 3). Together the class describe what they see using the strategy of *"I notice..., I wonder...."* I ask students to work in groups and write down their statements on a piece of A3 paper.

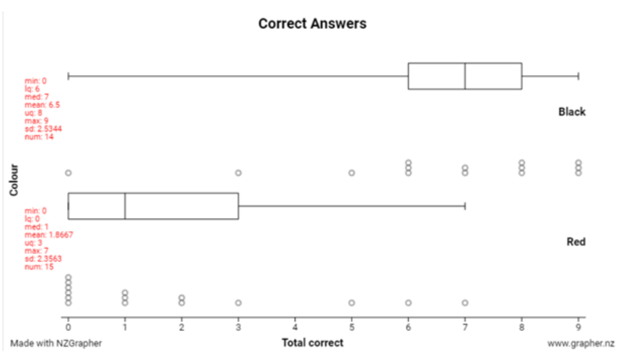


Figure 3: Boxplots, dot plots and summary statistics of correct answers. Black is the treatment group and Red is the control group

Students at this stage informally conclude that the teacher's training seems to have had an effect on the number of correct responses.

Student responses include statements like these:

I notice...

- Blacks did better than reds.
- Treatment group tended to have higher scores than the control group.
- **Treatment seems to have worked**.
- Black group are shifted further up the scale than the red group.
- 50% of black did better than 100% of red.
- Middle 50% of the red group are more spread out than the middle 50% of the black group.
- Red cluster at 0.

I wonder...

- If the black and red both had treatments, would they have gotten the same results?
- If **more time** would give different results.
- If **more participants** would result in a different outcome?
- If with easier times tables more reds might have higher scores?
- What the scores would be on a **younger age group**?
- What would happen if there was an **uneven number** of colours?
- If this is because of **academic ability**?
- If this was because all blacks had prior knowledge of a trick to answer better?
- If anyone from reds **knew the trick already**?
- If we would get the same results if red got treatment and black didn't (**so treatment and control swapped**)?
- If how some of the blacks scored lower than some of the reds?
- If we can make a call?
- If it **represents the population**?
- Would it be the same if we did it at a **different time**? Not first class back after the holidays?
- If we used a **different sample**, would it be the same?

Session Two

4. Critiquing the experiment

In session two we discuss what students noticed in the results. Students identify the "story" behind certain results.

> For example:
> "I'm the person who was in the black group and got 0, I couldn't follow what you were saying about the trick sorry."
> "Wow Emma was in the red group and she got 7 correct without any help"

I then facilitate further discussion on all the other factors that could have influenced these results.

> Student responses include statements like:
> "Some students can still remember how to perform traditional times table calculations while others can't."
> "Some students get stressed when tests are timed."

We discuss why we need random allocation of the experimental units to the treatment group. This leads to a discussion of how random allocation should balance out all the other factors that could explain the difference in the results between the treatment and control groups. These other factors are the "chance only" explanations for a difference between the results (Pfannkuch et al., 2013, 2015).

> I saw students thinking about "chance explanations" in statements like:
> "It could be that the trick didn't work but all the students who are good at maths just happened to be in the black group"
> "We need to use random allocation to have a mixture of abilities in each group."

Our reflections led us to wondering how we will know if the difference between the results is large enough to be sure it was not just due to luck or explained by any of the other factors that we identified. We discuss how what we see in Figure 3 includes both the treatment effect and chance, that is, some of our chance explanations. Specifically, the treatment effect in the Black group is made up of the treatment and chance. We then put our data into the *iNZight VIT Online Randomisation Test* tool (https://www.stat.auckland.ac.nz/~wild/VITonline/)) to help us to see the effect of re-randomising or reallocating our data to two different groups and the size of the difference between the medians that is likely just by luck or chance alone. (The *iNZight VIT Online Randomisation Variation* tool can be used as an alternative method to visually see the difference between the medians when the data are randomly allocated to two groups.)

5. Introduction to experimental design principles

At this point we discuss aspects of experimental design including:

- In an experiment everything must be exactly the same for both groups except the treatment
- Experimental units
- Random allocation of treatment to experimental units
- Controlling other sources of variation

We reflected on how our experiment ensured that both groups were the same and all potential sources of variation were controlled. We noted that the test was conducted by the same person and that all participants did the same test in the same room on paper. I prompt them to deliberate on issues such as, *What would happen if one group did the test online, the other on paper?* We consider how differences in mathematical ability are reduced via random allocation to two groups. Finally, we discuss how each of the factors that students commented on in their *I wonder …* written statements in session one were controlled.

An example of how we summarise our discussion on experimental design principles is shown in Table 1. It is important from a pedagogical perspective that the summary is co-constructed with the students at the end of the session two **after** they have participated in the experiment.

Table 1: Summary of the key features of our experiment

Design feature	Summary for our experiment
What type of experiment was this? Comparison of two independent groups design - what does this mean?	Each experimental unit (student) is either in the control or treatment groups. They cannot be in both groups.
Experiment units - who was the experiment carried out with?	The experimental units are the students in the year 13 Statistics class.
Response variable - what was measured for the experiment (what you hope will change as a result of your intervention)?	The response variable is the number of correct answers (out of 10) that each student achieves.
Treatment variable - what was the one thing changed between the two groups and how was it changed? What were the two treatment levels (describe in detail)	The treatment variable is whether or not the experimental units (students) were shown the 11s trick. The two treatment levels are: Shown Trick and Not Shown Trick.
Groups created using random allocation - why do we need to do this?	We randomly allocate the experimental units to the treatment groups to balance out all the variables we cannot control such as whether students are good at timed tests or feeling tired that day. These are all the chance explanations that could explain the difference in response variable (no. of correct answers).
Managing other sources of variation - what did the teacher/experimenter do to keep conditions the same for both groups? What other sources of variation (related variables) could have affected the results? Explain how they could affect the results. Describe how you could control some of these variables for the experiment and identify the ones you cannot control.	The teacher/experimenter has controlled a number of variables by administering the test to both the treatment and the control group at the same time, with the same instructions. Some controlled variables include: time of day, time allowed (2 minutes), comfort (temperature) and lighting in the room. Other variables that cannot be controlled include how tired the students were, whether they felt anxious about doing a test or not, whether they learned a new trick quickly or not, whether they could remember how to multiply by 11.
Describe how the experiment was carried out. What procedures did the experimenter follow? How was the data collected? Using what?	Following random allocation of the experimental units to two groups and random allocation of which will be the treatment group, the treatment group is shown the 11s trick while the control group is not. All students then complete as many multiplication questions with 11 as they can in 2 minutes. Students mark the test papers and the experimenter records the number of correct answers in a spreadsheet.

Subsequent lessons include use of the randomisation test and how to interpret the tail proportion.

The experiment gives an opportunity to include some reflection on maths anxiety and the potential issues with giving students times table tests in primary school. Year 13 students relate to these issues and find the brain-based learning ideas interesting (Boaler, 2019). The discussion fuels some ideas for experiments that they could design and conduct themselves. I include clips like: https://www.youtube.com/watch?v=HqultEUttE8 and https://vimeo.com/103921131.

Teacher notes

Students often transfer ideas about sample-to-population inference into this lesson as can be seen in their *I wonder* comments. Concerns about sample size and the sample adequately representing the population are evident. There was also discussion about fair testing with ideas being transferred over from their science classes. Each of these issues needed to be given time to consider in class.

To address students' concerns about sample size, we looked at the results from another class to see if our findings could be *replicated.* We discussed how replication of results across groups of similar sample size can help to establish a treatment effect. As well, it is useful to discuss how the treatment effect (the signal) would need to be large to be detected amongst the noise (chance variation) if the number of experimental units in each group were small.

With regard to the sample adequately representing the population, we discuss that in experiments the participants are not a random sample from the population, rather they are volunteers or a specific group. Therefore we can only draw a conclusion about the people in the experiment or about people very similar to those in the experiment.

The experiment implicitly follows the PPDAC cycle of enquiry. The PPDAC structure is overtly discussed in subsequent lessons.

References

Boaler, J. (2019). *Limitless mind: Learn, lead and live without barriers.* New York: HarperOne.

Pfannkuch, M., Budgett, S., & Arnold, P. (2015). Experiment-to-causation inference: Understanding causality in a probabilistic setting. In A. Zieffler & E. Fry (Eds.), *Reasoning about uncertainty: Learning and teaching Informal Inferential Reasoning* (pp. 95–127). Minneapolis, MN: Catalyst Press.

Pfannkuch, M., Wild, C.J., & Regan, M. (2013). Students' difficulties in practicing computer-supported statistical inference: Some hypothetical generalizations from a study. T. Wassong, D. Frischemeier, P. Fischer, R. Hochmuth, & P. Bender (Eds.), *Mit Werkzeugen Mathematik und Stochastik lernen [Using Tools for Learning Mathematics and Statistics]*, (pp. 393–403). Wiesbaden, Germany: Springer Spektrum. doi: 10.1007/978-3-658-03104-6

Materials required

Playing cards, printed test sheets, countdown timer (cell phone), A3 sheets
NZGrapher or *iNZight Lite*
VIT Online

Copyright information