VIRTUAL EXPERIMENTS TO TEACH EXPERIMENTAL DESIGN: A WEB-BASED TOOL FOR BIOSTATISTICS STUDENTS BRIDGING THE GAP BETWEEN DATA COLLECTION AND STATISTICAL ANALYSIS

Charlotte M. Jones-Todd and Amy Renelle
Department of Statistics, University of Auckland, Auckland, New Zealand
c.jonestodd@auckland.ac.nz

Biostatistics students expect to graduate with the required tools and knowledge to collect and analyse their own data. Many go on to careers that require them to design their own lab- or field-based experiments. However, the examples introduced throughout their courses are textbook: data have already been collected appropriately, wrangled, and cleaned. Dealing with pre-provided, historic data does little to motivate student engagement with the underlying statistical concepts and fails to bridge the gap between statistical theory and application. Here, we discuss an application that provides students with the experience of designing an experiment and collecting their own unique dataset in a cost- and hassle-free way.

INTRODUCTION

Using real-world examples contributes to creating an active learning environment, consistent with the constructivism learning paradigm (Mcleod, 2019; Miller, 2002). Students often find textbook scenarios and real-world examples difficult to connect, so the integration of real data in statistics classrooms, as promoted by Carver and colleagues (GAISE College Report ASA Revision Committee, 2016), is important to consider in the development of activities. However, real-world data can be messy, and real-world relevant data collection can be too unpredictable to reliably generate data that highlight the specific statistical concepts being taught (Halley, 1991; Schneiter et al., 2018). Dealing with raw, messy, data can be beneficial for immersing students in apprenticeship-style learning, but it can also distract from the key concepts that are the focus of the task "*Poor response rates, coding errors, and sampling errors are absent from simulated data; thus, students and instructors receive a ready supply of 'clean' data for demonstrating statistical concepts and principles*" (Halley, 1991, p. 518).

University biostatistics courses are constantly balancing teaching statistical theory and practice. These courses are designed to give students the knowledge and tools to progress in their careers as field or lab scientists (or otherwise). However, they cannot focus on statistical theory as well as the time-consuming procedures that are involved in data collection. A purely statistical experimental design course may focus on the more mathematical aspects of the designs (e.g., orthogonality and constraints, etc.). A more applied course will likely focus on data collection (e.g., the mechanics of replications, etc.). However, experimental design requires an understanding of both as well as forethought as to what analyses will be appropriate.

Simulated student-data is a possible alternative that still provides an active learning environment (Miller, 2002). Simulation allows for students to actively engage in the whole experimental design process but circumvents the time required to carry out the experiment (Ketelaere et al., 2014). Situated within a real-world context, *"... simulation can be a way to make sense of the world ..."* (Burrill, 2021, p. 5) while focusing on the key ideas at hand. A virtual setting is a cost- and hassle-free way to emulate a real-world experiment whilst giving students ownership of their data and providing a rigorous scaffolding enabling each example to be tailored to highlight the desired learning outcomes. Specifically designed statistical software tools can provide students with a virtual, explorative learning environment and interactive tasks that promote participation (Burckhardt et al., 2018; Erickson, 2013; Kaplan, 2014; Sedlmeier, 2014; Wild, 2018). From automating arduous processes to simulating many observations efficiently, there is a wealth of literature demonstrating the benefits of using digital technology to teach statistics.

The *vested* application (Jones-Todd , 2021) provides students with the experience of designing an experiment and collecting their own unique dataset. This software offers students a classroom-friendly, hands-on experience of data collection, and its small-scale tailored framework is simple to implement from a lecturer's perspective. In this paper we outline the virtual experiments that the application offers and discuss the learning objectives targeted.

THE APPLICATION

The software implements three virtual experiments, each with varying degrees of complexity and scenarios to consider. Alongside focused exercises, the virtual platform facilitates student discussion around the principles of experimental design directly linking them to real-world applications students may face later in their careers. This application is currently hosted online and was built using the statistical language R (R Core Team, 2021) and the R package Shiny (Chang, et al., 2021). All source code for the application is open-source (Jones-Todd & Jones, 2020).

The virtual experiments currently implemented are detailed below. In each case arbitrary units are discussed; this is by design so that students are encouraged to think about effect sizes and relative changes.

The Country Pumpkin Experiment

This experiment is based on the example outlined in Pollock, Ross-Parker, and Mead (1979). Students are asked to investigate pumpkin growth. They are faced with 16 plots in a greenhouse arranged so that eight are along the south-facing wall and eight along the north-facing wall. A control panel enables students to choose the variety of pumpkins to plant in each plot (either Buttercup or Kumi Kumi). In addition, students can choose whether the plots are subjected to supplemental heat and controlled light. They are asked to ascertain what combinations of these treatments lead to the maximum possible pumpkin yield. They are also told that the plot orientation within the greenhouse may have an effect.

After students have set up their experiment, they can download their data in a CSV file for analysis in their software of choice. These data are simulated along with some random noise following the relationships shown in the left-hand plots of Figure 1. Here, we can see that there is generally little difference between varieties. Overall, supplemented heat increases yield, but controlled light has only a small effect. Generally, the effect of orientation is the same across varieties in each combination of heat and light treatments: south-facing plots induce higher yields on average. However, there is an interaction effect simulated so that a Kumi pumpkin subjected to natural heat and controlled light in a south-facing plot has a much higher yield on average than the Buttercup.

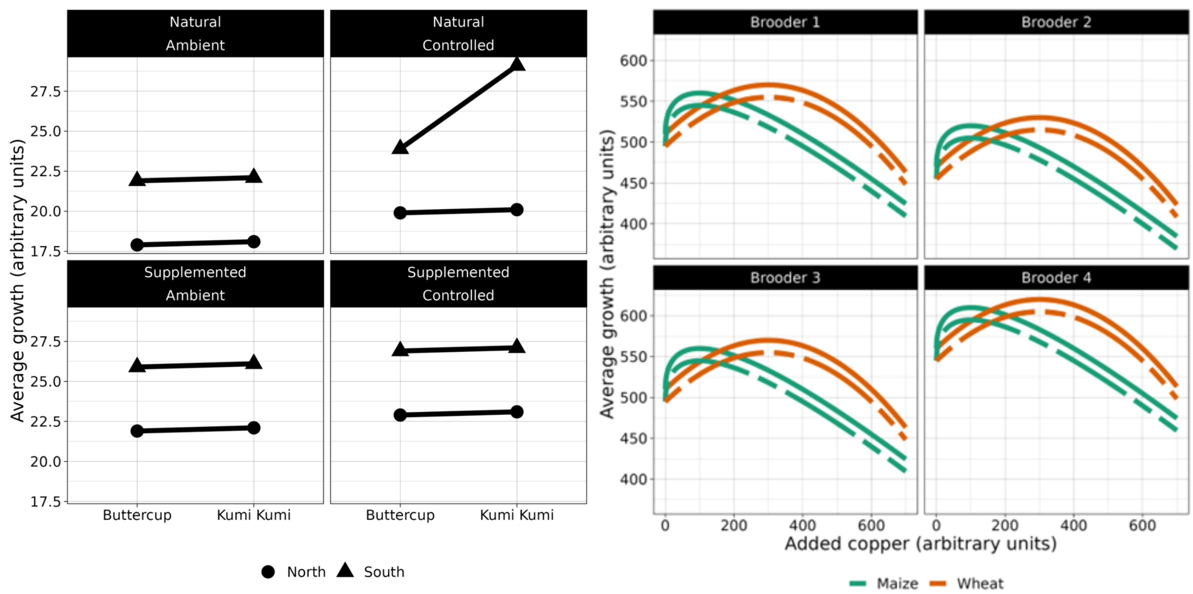

Figure 1. Left: simulated average pumpkin growth depending on each treatment level of variety, heat, and light. In addition, the effect due to plot orientation is shown. Right: Simulated relationship between pullet growth and copper depending on each of the two diets (maize and wheat). The relationship varies slightly depending on the brooder (shown in each panel); there is also a small, simulated tier effect, which is not illustrated here.

The Chook Eggsperiment

This experiment is based on an example outlined in Mead (1973). Students are asked to ascertain what optimum amount of copper added to a chick's diet will maximise the growth rate. They are told that 150 units of copper added to either a wheat or a maize diet has been found to improve growth. There is also known to be a value beyond which the amount of copper added becomes toxic. They are not told what this might be. Students are faced with a control panel and a schematic of cages. In total, they have 32 cages each with 16 chicks to feed. Each set of eight cages is known to be from one of four possible brooders, with their cages arranged in one of four tiers.

Students can choose either a wheat or a maize-based diet for each cage and choose the amount of copper to add (between 0 and 700 units). This procedure is repeated for all four brooders. Again, after students have set up their experiment, they can download their data in a CSV file. This is simulated along with some random noise following the relationships shown in the right-hand plots of Figure 1. The relationship between copper and growth is non-linear, the optimum amount for a wheat-based diet is around 300 units, whereas for a maize-based diet the optimum amount is around 150 units. There is a brooder effect shown in each panel of the right-hand plots of Figure 1 as well as a slight effect due to the tier of the cage.

The Head Tomato Experiment

This experiment is based on that outlined in Darius, Portier, and Schrevens (2007) where students are asked to determine the optimal dose of manure to add to a fertiliser recipe to maximise tomato yield. Students are faced with a schematic of a greenhouse floor (of 6 x 6 cells) where they are asked to drag and drop tomato seedlings to grow. They are told that the heat and light intensity varies across this space, although they are not told how. Tomato seedlings (labelled by their initial height) can be chosen from six different trays and placed anywhere on the greenhouse floor. If a student wishes to discard a seedling there is also an option to do so. Students do not have to use every seedling. In addition, multiple seedlings may be placed in each cell.

Two other control panels allow students to create four fertilizer recipes, each with a user-chosen proportion of manure to add and dictate which seedling receives which fertilizer. Finally, the students should set the number of days they wish to grow the seedlings (with a set maximum). After everything is set up the data can, again, be downloaded in a CSV file. The simulated relationship without the added random noise is shown in Figure 2. The relationship between manure proportion and growth is non-linear, this changes as an additive constant for the number of days the plants are left to grow. In addition, the heat and light effects for each cell are shown in the left and centre plots, respectively, of Figure 2. Here, the darker red indicates a greater effect on average growth.

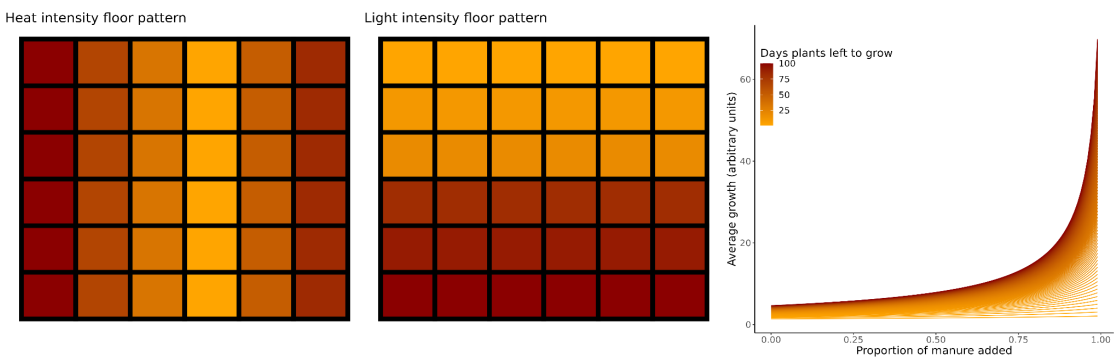

Figure 2. Left and centre plots, respectively, show the simulated heat and light effects for each of the 6 x 6 cells. The right-hand plot shows the simulated relationship between tomato growth and manure added to the fertilizer recipes; the colour scale illustrates how this relationship changes depending on the number of days the plants are left to grow.

LEARNING OBJECTIVES

In this section, we discuss the targeted learning objectives and how they link to statistical techniques in analysing these data appropriately. Each virtual experiment offers a different level of complexity where the number of treatments to the extraneous factors involved differ. Each affects the outcome, and each should be considered in the design of the experiment.

Encouraging students to think of the different potential sources of variation whilst designing the experiment reinforces the effects students should also account for when they are asked to analyse their data. Listed below are some typical learning objectives associated with teaching experimental design.

1. Identify the experimental unit(s) and the observational unit(s).
2. Identify and list the treatment(s) and the treatment levels.
3. List and describe the three main principles of experimental design. Discuss the role they play in a given experiment.
4. Define a fixed and random effect in the context of experimental design.
5. Carry out appropriate analysis for given experimental data (changes depending on stage, e.g., from linear regression to a mixed effect model to a non-linear mixed effect model, etc.).
6. Communicate statistical concepts and experimental outcomes clearly using language appropriate for both a scientific and non-scientific audience.

Figure 3 below shows the end-to-end progression for the Chook eggsperiment. Each section of the design and setup can be linked to the objectives listed above. The experimental design ultimately leads to data analysis, which can take a myriad of techniques from visualisation to non-linear mixed effect models. The vested application enables this aspect to be molded depending on the level of the course.

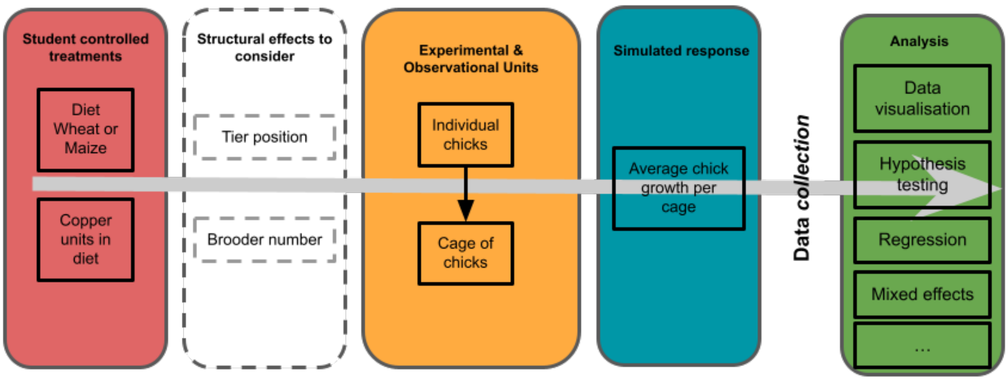

Figure 3. A flow diagram showing the end-to-end progression for the Chook eggsperiment. Similar schema can be constructed for both other vested experiments. From controlled inputs and structural effects students are guided through a typical design-based framework and rewarded with direct access to their data so that they can immediately investigate their hypotheses.

In-Class Use

The application has been used in class twice this semester (semester 1, 2022). In an undergraduate class (University of Auckland, 2021b) of approximately 300 students and in a postgraduate class (University of Auckland, 2021a) of 18 students. In each instance, students are given one of the scenarios and are asked to design their experiment given only the background information written at the top of each page. They are then asked to write up their design strategy after being prompted to specifically mention any sources of variation they had to consider. Additionally, they are asked how they, if at all, controlled for these (e.g., via randomisation, replication, or blocking).

As a stepping-stone to analysis, students are asked to download the data and produce a plot that they deem best to illustrate any relationships or effects that may exist. Finally, having built up from the design strategy, students are asked to analyse their data using any of the techniques that have been covered in the course. This activity is designed to target the other listed learning objectives (2, 4, and 5) in addition to those listed in the previous section.

The vested virtual setting facilitates this framework enabling students to start with data collection following this through to the statistical analysis all within a short time period. There are no monetary and very few time-related costs involved. The online setting means students can restart and rerun their experiments as many times as they wish. Thus, students are given the opportunity to trial and error different scenarios, empowering them to explore the (simulated) mechanisms driving the scenarios they are faced with.

- "*I really liked using the [software], it is really creative and a fun, interactive way to set up an experiment and collect data*."—Anonymous student.
- "*Really easy to use the application to produce the data, ... The pumpkin scenario was also very easy to follow and I didn't need any background knowledge to understand the experiment."*—Anonymous student.

DISCUSSION

Exposure to real data collection, as recommended in the GAISE College Report (ASA Revision Committee, 2016), is often not practical in a higher education environment, especially when teaching semesters may be relatively short (typically 14–15 weeks including breaks) compared to the time required to conduct and monitor experiments (e.g., 14 weeks for the Head Tomato experiment). The vested application allows students to have a similar level of engagement to the experimental design processes and decision making as they would have in practical settings. The end-to-end process, from design to analysis, creates a seamless stream for students to follow through from their initial design discussions to the final data analysis.

In addition, a virtual setting gives the benefit of live feedback and allows users to see the impacts of variation almost immediately. That is, running a virtual experiment allows the students to think about the impact their experimental design choices have and assess them immediately without the time and monetary constraints of real-life experiments.

To date, vested has been used in both undergraduate and postgraduate biostatistics courses at the University of Auckland. However, it, alongside targeted exercises, could easily be adapted for different learning environments (e.g., secondary schools). Additionally, there is scope for adding additional virtual experiments, which can be constructed to mirror the "real-life" scenarios students may face in their academic careers (e.g., microbiology, medical trials).

ACKNOWLEDGEMENTS

With thanks to Sophie Jones for her help with bug fixing as well as for her help, alongside Marina Klemm and Daniel Harrison, running a student focus group to trial the application before going "live" in class. Thanks also go to the 16 student participants of the trial focus group.

REFERENCES

Burckhardt, P., Chouldechova, A., & Nugent, R. (2018). TEACHIT: turning the classroom into a research laboratory via interactive e-learning tools. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* ISI/IASE. https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_9J3.pdf?1531364302

Burrill, G. (2021). Simulation as a tool to make sense of the world. In R. Helenius & E. Falck (Ed.), *Statistics education in the era of data science. Proceedings of the Satellite conference of the International Association for Statistical Education.* IASE. https://doi.org/10.52041/iase.oysdd

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., . . . Borges, B. (2021). *Shiny: Web Application Framework for R.* (R package version 1.7.1.) [Computer software]. https://CRAN.R-project.org/package=shiny

Darius, P. L., Portier, K. M., & Schrevens, E. (2007). Virtual experiments and their use in teaching experimental design. *International Statistical Review / Revue Internationale de Statistique, 75*(3), 281–294. https://doi.org/ 10.1111/j.1751-5823.2007.00028.x

Erickson, T. (2013). Designing games for understanding in a data analysis environment. *Technology Innovations in Statistics Education, 7*(2), 1–13. https://escholarship.org/uc/item/31t469kg

GAISE College Report ASA Revision Committee. (2016). *Guidelines for assessment and instruction in statistics education (GAISE) college report 2016.* American Statistical Association. http://www.amstat.org/education/gaise

Halley, F. S. (1991). Teaching social statistics with simulated data. *Teaching Sociology, 19*(4), 518–525. https://doi.org/10.2307/1317899

Jones-Todd, C. M. (2021). *vested. Virtual experiments to teach experimental design* [Computer software]. https://statbiscuit.shinyapps.io/vested/

Jones-Todd, C. M., & Jones, S. (2020). statbiscuit/vested (ICOTS release) [Computer software source code]. https://doi.org/10.5281/zenodo.6795867

Kaplan, J. (2014). Clickers, simulations, and conceptual understanding of statistical inference. In K. Makar, B. de Sousa, & R. Gould (Ed.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS 9, July, 2014).* ISI/IASE. https://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_4B1_KAPLAN.pdf?1405041622

Ketelaere, B. D., Siebertz, K., van Bebber, D., & Rutten, K. (2014). The garden sprinkler: An interactive web-based application for teaching design of experiments. In K. Makar, B. de Sousa, & R. Gould (Ed.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics.* ISI/IASE. https://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_C266_RUTTEN.pdf?1405041919

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education, 2*(1), 1–37. https://escholarship.org/uc/item/38p7c94v

Mcleod, S. (2019). *Constructivism as a theory for teaching and learning*. Simply Psychology. https://www.simplypsychology.org/constructivism.html

Mead, R., & Freeman, K. (1973). An experiment game. *Journal of The Royal Statistical Society Series C, Applied Statistics, 22*, 1–6.

Miller, J. B. (2002). Examining the interplay between constructivism and different learning styles. In B. Phillips (Ed.), *Developing a statistically literate society. Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6).* https://iase-web.org/documents/papers/icots6/8a4_mill.pdf?1402524963

Pfannkuch, M., & Brown, C. M. (1996). Building on and challenging students' intuitions about probability: can we improve undergraduate learning? *Journal of Statistics Education, 4*(1), 1–15. https://doi.org/10.1080/10691898.1996.11910502

Pollock, K. H., Ross-Parker, H. M., & Mead, R. (1979). A Sequence of games useful in teaching experimental design to agriculture students. *The American Statistician, 33*(2), 70-76. https://doi.org/10.1080/00031305.1979.10482663

R Core Team. (2021). *R: A language and environment for statistical computing. R Foundation for Statistical Computing.* Vienna, Austria. https://www.R-project.org/

Schneiter, K., Christensen, L., & Lee, V. R. (2018). Using personal activity data in an undergraduate statistics course. In M. A. Sorto, A. White, & L. Guyot (Ed.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* ISI/IASE. https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_9I1.pdf?1531364301

Sedlmeier, P. (2014). Technology for developing statistical thinking: A psychological perspective. In K. Makar, B. de Sousa, & R. Gould (Ed.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014).* https://icots.info/9/proceedings/pdfs/ICOTS9_9F2_SEDLMEIER.pdf

University of Auckland. (2021a). *Advanced Biological Data Analysis*. BIOSCI738. https://courseoutline.auckland.ac.nz/dco/course/BIOSCI/738/1213

University of Auckland. (2021b). *Quantitative Biology*. BIOSCI220. https://courseoutline.auckland.ac.nz/dco/course/BIOSCI/220/1213

Wild, C. (2018). Gaining iNZights from data. In M. A. Sorto, A. White, & L. Guyot (Ed.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* ISI/IASE. https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_9A3.pdf?1531364299