Case Studies in Research Methods Pedagogy

Teaching computational statistics through active learning

By Sarah Lewthwaite, University of Southampton

This case draws upon ten hours of observation, field-notes, teaching materials, one teacher and two student interviews, student notes and informal research conversations with students across the sessions. The narrative seeks to demonstrate and situate the detail of a singular teaching approach: active learning. This is done to draw out the teacher's craft knowledge – the strategies and tactics that make the teaching come alive. Many agree that teaching quantitative methods involves active learning [1, 2, 3] in which learners conduct exercises; in this case, we appreciate the minutiae of the work and the sheer graft that the teacher and learners put in. In this instance, to make his approach work, the teacher, who we shall call Kevin, has invested considerable strategic work in preparing his content, and he employs great tactical skills in his exposition in class. An active learning approach was consistent across the pedagogy specified and enacted by the teacher, and experienced by the learners. Such consistency across these pedagogic dimensions is by no means inevitable [1].

We have convened in a small, tiered lecture theatre. Students with laptops and bottles of water are dotted across rows of fixed desk. At the front there are whiteboards, a large data projector screen and lectern. Kevin, an established teacher/ researcher who is seemingly immune to jetlag, takes this class.

The classes are 'all meant to be very hands-on' Kevin begins as he addresses the class: 'I'll lecture, and through the lecture I'll pull up software and we'll walk through the computer code and what the output will be', then for the last hour the students will work on exercises. 'So if you go to GitHub you'll find student materials, the slides, the code, exercise and exercise datasets'. 'There's no grades' he adds. 'Then' he explains 'for the last 10 minutes, I'll walk through the answer key. So you have a sense of me doing it, but also get to work through it yourself. [You'll] Essentially repeat what I did in class, but with a new data set. All sessions will follow that basic pattern. I'll be here, as you work through, hands-on and I can help you. We'll take a short break at around 10.30.'

Now Kevin begins to lecture, taking the group through a specific method operation. The whole group is attending closely. Next to me, a student, Gerhard, begins writing furiously. Soon, Kevin moves from his slides to clean the white board, and then marks out a matrix to expand upon the slides. The lecture moves easily from the minutiae of the equations, through different types of applied and theoretical examples.

Kevin's examples are related to debates, the occurrence and avoidance of controversies, back and forth, to 'why' this approach and 'what' key applied issues are. He mentions the history of the statisticians who developed the methods, their rivalries and friendships.

Teaching approach

Kevin tells me his approach is 'active learning'. While lecturing has value, Kevin states 'I don't think they're [students] really going to pick it up until they do hands-on [in class] homework with data'. The class is structured to this end. Kevin's strategies, tactics and tasks are all orientated to this principle.

Teaching tactics

Kevin's exposition is also peppered with pedagogic rhetoric, 'typically', 'note that', 'remember' and so forth. This talk is *tactical*, drawing attention to nuance regarding his code, the equations and overarching content.

Prepared examples are deployed fluently from multiple disciplines, in concrete terms (code from actual experiments from multiple disciplines, medicine and education) alongside more spontaneous context: 'social psychology has been criticised a lot recently…', 'if you've done econometrics, you've probably seen this already'. He identifies whether something (a framework, a model) is rare, the limits, challenges and benefits of using a given method in a particular way, as well as giving some information that gestures to the status of method, how it's commonly used and where.

Kevin also assails impromptu points of interests, for example, what journals are looking for or what content is untouched by contemporary textbooks. He flips back and forth within the slides to cast insight on use of specific equations in light of a new scenario, or alternative/additional application.

Kevin moves to the light switch and checks, 'Is that highly visible to everyone in the back?' The class is nodding. With his laptop broadcasting its screen through the projector, we now see Kevin switching between code libraries and statistical package R. There is code for the students to manipulate.

He begins working with R. 'So I'm going to run a little loop - you don't need to know R programming because I've got a little bit of code you can use later...'. In this part of the class, the depth of preparation becomes visible:

> 'So let's take a look, I've got data for that too'.
>
> 'For completeness, I wrote the whole thing out … So I coded it two ways'
>
> 'again, this isn't how you would usually do it, this is an instructive example'

Robyn, a student (and an associate professor and methods teacher, herself) whispers to me 'this is the gold standard'. Later we talk and she explains:

> Kevin's structure [is] the gold standard … everywhere else I've seen this, it's been the most effective. You want to start off with framing it, right? And then you present the theoretical specification, and you move back and forth between real world examples, so to contextualise it, right? And then you come with a problem that's using software, okay. And then you give students the opportunity to practice, and kind of alter [it] a little bit, so it's, so now you're testing their range of thinking … I think that's really, really effective, but that kind of pedagogy would not be effective if you are not organised.

STRATEGIC USE of EXAMPLES

The presentation and interaction with the data is carefully ordered and planned through strategically chosen examples: 'The first group we can learn about ...', 'I'll show you here, what we're going to do…', '[this] is a great example, it's got [x, y, z] it's got everything'.

STRATEGIC USE of RESOURCES

Mirroring the move between (strategic) prepared content and more tactical delivery, Kevin's interaction with different learning resources is notable. Slides are heavy with tables, equations, bullet points and occasional graphs. Kevin then diverts to the whiteboard to provide additional visualisations. He is focussing on teaching through data and code, and the pedagogic rhetoric reflects this. He illustrates - 'You can see in the data'; tests - 'what has to be true here?'; evaluates - 'but what's wrong with this..?'; modelling how value is placed. As the exposition moves to demonstration, Kevin's tactical pedagogic repertoire adapts accordingly.

Organisation is not the sum of the challenge, however, as Kevin explains:

> the other flipside is, it's actually easier to get the students to the point where they can spit something out of a computer, [but] statistics is only relevant if you can be very clear and interpret what comes out of that computer, because computers … they're very smart and very dumb, they can spit out pure nonsense.
>
> But I think the [in-class] homework is where that, … And it's easier to say 'well let's go and find a good dataset and maybe we'll just skip this one'. So to me actually one of the more important pedagogical aspects is, I mean I have to be clear in class, but then what is more important is actually setting up a set of exercises for every topic, and making sure they've worked through it.

As Kevin starts to 'Walk Through' the code, he begins to think aloud, taking the class through his reasoning and concrete actions, whilst gesturing to the statistical rules and coding grammar that he's adhering to. R Studio mediates this activity both as a frame and as a tool. Kevin locates his actions both conceptually: 'Basic principles are always the same …' whilst highlighting technical skill, 'the only way to do this really is with a computer … so we'll let the computer take care of this'. The two facets of methods competency, theory and technique, are clearly applied and specified: 'This is one of those things that is very easy to see on a computer but less obvious in the framework'.

Kevin is moving swiftly between his slides, live code, the whiteboard and the problems set out in the worksheet. As a researcher, taking photos and making recordings alongside a mess of notes is challenging, so it must be so for the students. Gerhard is still notetaking at pace; 'when I write it down I remember better' he says. Christa finds this is very advanced, 'it's hard to keep up' she reflects mid-course, adding 'I can because I have some background knowledge.'

Once the walk-through is complete and outputs demonstrated, Kevin turns to the in-class homework: 'I personally think this kind of exercise is very interesting'. Exercise IV is brought up on the projector. 'There are some nice subtle things going on there', Kevin continues. 'So let's have a crack at that. I'll be here to answer questions'. Those without laptops begin to negotiate with their neighbours - there is some murmuring around the room. The students are pulling up GitHub and opening the files and examples on their screens.

Kjeld helps Gerhard to find the right files, 'Did you download the right one?', 'I downloaded it yesterday, so...'

Kevin directs, 'all the data is in Stata but you can redirect into R. You guys know how to do it?' Kevin takes them through how to add to R, 'I tend to use Stata data files and then move into R, so I can keep clean data files elsewhere.' He clarifies which files are necessary to the worksheet task. The students' progress with Kevin's 'homework in class'.

PEDAGOGIC CHALLENGE

At interview, Kevin speaks of the 'load on the instructor' directly. 'Mining datasets, writing the exercises and writing the answer key' are some of the 'most challenging aspects of teaching statistics'. 'That can be very time-consuming, and it's easy to just say, "No, let's not have a homework on this topic", but that's actually 90% of, not just them doing it, but then after they do it, stepping back and reviewing every aspect of it post-[exercise].'

STUDENT STRATEGIES

Some students had active strategies for managing their learning, particularly in terms of multi-modal note-taking. Robyn reflects: 'the important things he [Kevin] would say that are not in the notes or in the R student files … so for the lecture, I get the slides ahead of time. I have to be organised, and I need to organise the files the way that I learn …So I took his [Kevin's] organisation and I adapted it. For example, I like all my R files in one folder… so that when I pull them out-up in R, when I set my working directories, all coming from one … next for the lectures, I have PDF files, and then I just use the Note feature in PDF to write notes next to something. Like say if he said, "[X] is [X] whether they'd assigned a treatment or not", right?

Alana, Kjeld, Oscar, Nina and Gerhard lean over their desks. Alana turns to Kjeld to ask how he got the file open. ‘Ahhh, thank you’. Then Oscar asks a question and along the row, Nina pitches in: ‘Have any of you guys calculated the means …?’. Gerhard responds, Nina clarifies a point, highlighting onscreen. Gerhard sings ‘you’re so smart!’.

Madeline, Teyo and Alana huddle around one screen.
Madeline: [gesturing to screen] Here, here. Can we change it? OK, cut and paste it.
Teyo: I don't know
Madeline: Do you think that'll work?
Teyo: Not really [laughs]
Madeline: maybe if we also put it [code] here ... NICE!
Teyo: [whispers] Perfect! [They look at the screen.]
Madeline: What does it mean?! [All three laugh] This is treated compliant, this is treated non-compliant […]
Alana: The assumption holds true …

All the students are now in R, some flip between the slides and the software. Kevin begins moving round the room, suggesting technical fixes as well as code ‘use an ampersand for “and” or it’ll give you an error’, and giving statistical support. As the class settles, leaning into screens with hunched shoulders, Kevin watches intently. For Kevin, this period is essential. It serves more than one strategic pedagogic end.

Time passes quietly except for the rapping of keyboards, the occasional question or whisper. One hour in Gerhard has successfully run his code. He gets up and leaves for fresh air. Ten minutes before the class ends, Kevin delivers the Answer Key, with further technical nous: backstage knowledge of what can crash R, which code libraries are most effective and so on. The course finishes to short applause.

STRATEGIC USE of ACTIVE LEARNING

At interview Kevin is clear, active. task-based learning can change the nature of the class, opening an essential questioning space.

‘a lot of the students won’t ask questions in class … so the last hour or so of the session is devoted to exercises that they work on, people tend to be more willing to ask questions because they’d ask questions when you come to them and you consult with them, one-on-one, then they can ask a question in a setting where not everybody hears. I’ve also noticed there are even people who are reticent to ask questions in that setting, they’re staring aimlessly at their screen. If you walk up and say, ‘does it all make sense?’, when prompted they’ll then ask a question, and then you can get them unstuck.

References

1. Allen, P. J. and Baughman, F. D. (2016) Active learning in research methods classes is associated with higher knowledge and confidence, though not evaluations or satisfaction. *Frontiers in Psychology. 7: 279.*
2. Baglin, J., Reece, J. & Baker, J. (2015) Virtualising the Quantitative Research Methods Course: An Island-Based Approach. *Statistics Education Research Journal.* 14(2): 28-52.
3. Lovekamp, W. E., Soboroff, S. D. & Gillepie, M. D. (2017) Engaging students in survey research projects across research methods and statistics courses. *Teaching Sociology.* 45(1): 65-72.

This case study is part of a series produced from the Pedagogy of Methodological Learning study. Look out for them on the NCRM website. We also appreciate feedback to inform future work.

National Centre for Research Methods
Social Sciences
University of Southampton
Southampton, SO17 1BJ
United Kingdom.

Web www.ncrm.ac.uk
Email info@ncrm.ac.uk
Tel +44 23 8059 4539
Twitter @NCRMUK

To cite this paper: Lewthwaite, S. (2021) Teaching computational statistics through active learning (Case Studies in Research Methods Pedagogy). National Centre for Research Methods. http://eprints.ncrm.ac.uk/4404/

See the complete case study series here: https://www.ncrm.ac.uk/resources/trainers/quickstart.php