# Refining a Workshop Intervention for Scientists on Inference and Scientific Methodology

Robbie Clark[1,2], James Ladyman[3], Marcus Munafò[1,2]

[1]School of Psychological Science, University of Bristol

[2]MRC Integrative Epidemiology Unit at the University of Bristol

[3]Department of Philosophy, University of Bristol

## Abstract

**Background:** Researchers often lack in-depth training in philosophy of science and epistemology. These subjects can help to foster a critical appraisal of research methodologies and inferential practices. Consequently, we created a workshop on inference and scientific methodology for researchers, which we ran twice with researchers at University of Bristol.

**Aims:** We aimed to learn (1) whether attendees find the workshop beneficial, and (2) how to improve the workshop.

**Methods:** We conducted two studies, using questionnaires before and after the workshops to answer our research aims. We ran the workshop once ($N$ = 7 participants), refined it with feedback, and then ran it again ($N$ = 5).

**Results:** Both studies suggested that the workshop can benefit researchers. Beginning with a visualisation of individual research practices, it can build an understanding of different types of inference and how they are used in practice. This can facilitate critical thought about research practices and the inferences made therein.

**Conclusions:** The workshop appears beneficial to pursue further. We suggest additional implementation and testing, collecting data on attendees' attitudes, knowledge, and behaviours. These should be assessed both immediately and several months after the workshop.

**Keywords:** *intervention development, workshop, philosophy of science, scientific method, inference, induction, deduction, abduction.*

## Introduction

Scientific training often involves comprehensive teaching of domain-specific methods. Scientists are taught *what* to do in order to conduct research within their field. Meanwhile, training in the underlying justification for these practices is often neglected (Grüne-Yanoff, 2014). Understanding *why* things are done in a certain way enables researchers to question and appraise the appropriateness of their practices. The need for this has been particularly apparent during the last decade within disciplines such as psychology. As questions have arisen concerning the quality and reliability of published research, people have returned to philosophical foundations in order to address the problems raised and to improve methods (Albert et al., 2020; Derksen, 2019; Fletcher, 2021; Flis, 2019; Morawski, 2019; Nosek et al., 2021; Reiter, 2013).

A greater focus on training scientists in philosophy of science and scientific methodology (i.e., the study of the scientific method) should improve researchers' understanding of the underlying rationale for their research practices (Grüne-Yanoff, 2014; Meehl, 1993). Evidence from a workshop intervention called the Toolbox Project has indicated that philosophical discussion on epistemology and ontology helps to foster more fruitful cross-disciplinary collaborations among researchers (Eigenbrode et al., 2007; O'Rourke & Crowley, 2013). Some of the benefit comes from the fact that such discourse provides a common vocabulary, which enables researchers from different disciplines to work together effectively. Moreover, it encourages attendees to consider their own ontological and epistemological assumptions, which are sometimes held implicitly and unchallenged.

Inspired in part by this idea, we developed an interactive workshop on inference and scientific methodology at the University of Bristol, which we ran twice. The workshop was designed for researchers at postgraduate level and beyond. Broadly, the workshop consisted of two sections, the first on scientific methodology, the second on inference. Initially, attendees discussed pictorial representations of what the entire research process looks like for them (see Figure 1.a. for an example of one attendee's *methodological schematic*). Afterwards, we discussed each of three types of inference in turn (induction, deduction, and abduction).

Attendees learned about each type of inference and how they are commonly used (we assumed no prior knowledge). We hoped to elucidate the interconnectedness and, more crucially, the fallibility of inferences. Ultimately, we wanted attendees to identify how they can arrive at more trustworthy conclusions, recognising where errors are most likely to occur, and the need for humility within science (Fjelland, 2022). Attendees consistently referred back to and annotated their methodological schematics, considering how they use the three types of inference throughout their research (see Figure 1.b. for a schematic at the end of Workshop 1). Thus, the methodological schematics served to make the abstract terminology of inference relatable and relevant. Moreover, the schematics helped attendees to visualise where they make the most crucial inferences, as well as the more easily forgotten assumptions.

**Figure 1.a.** A methodological schematic, photographed at the beginning of Workshop 1.

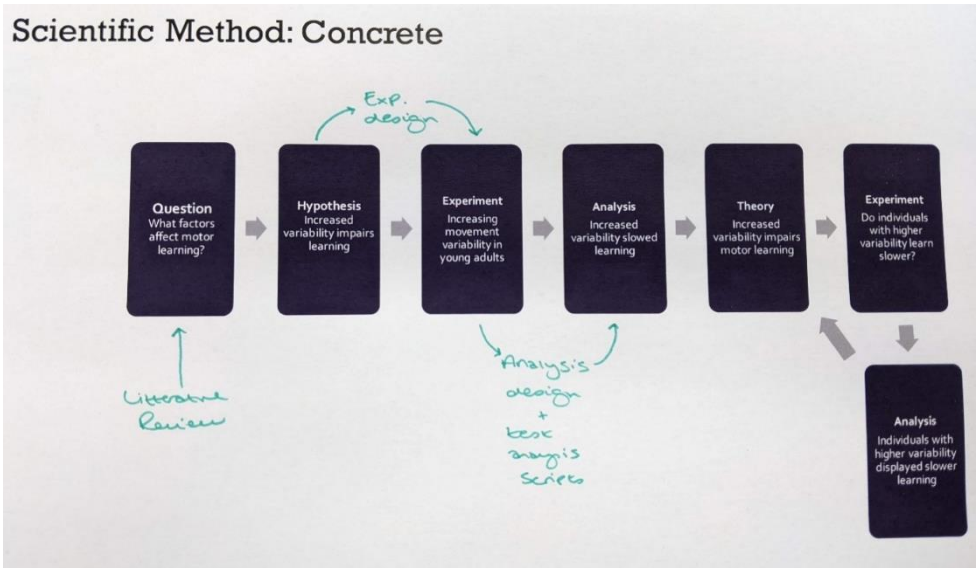
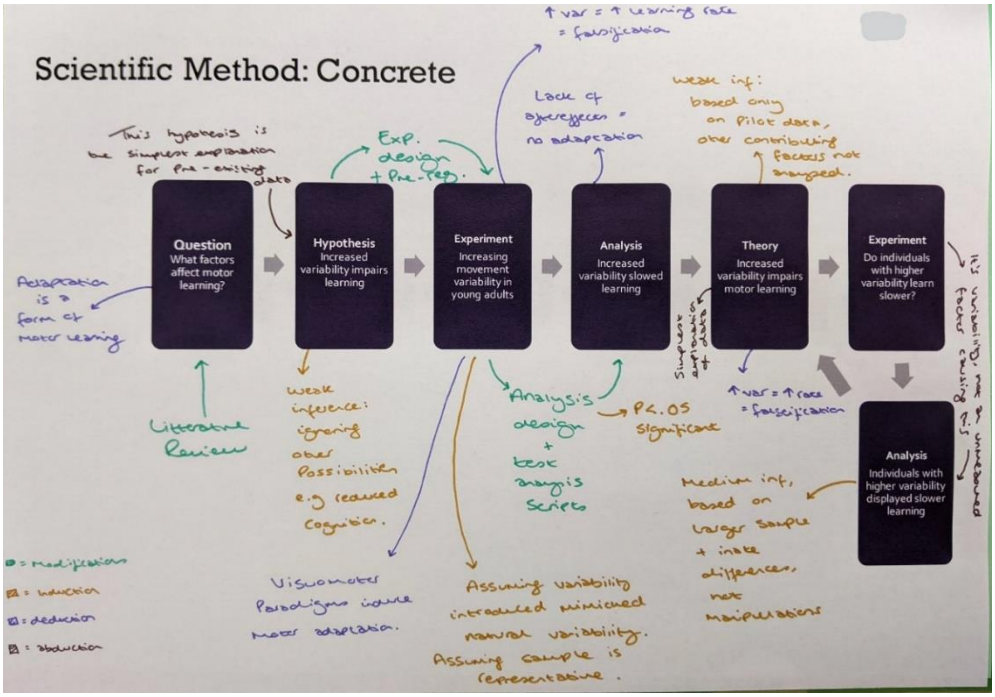



**Figure 1.b.** A methodological schematic, photographed at the end of Workshop 1.




*Attendees were encouraged to annotate their schematics during the workshop when identifying the key inferences that they make.*

We developed the workshop materials during 2021 and ran the workshop twice. After the first workshop (Study 1), we refined the materials and format before running it a second time (Study 2). Both studies involved the same two primary research questions (RQs). RQ1: whether attendees find the workshop beneficial. RQ2: how to maximally improve the workshop for future iterations. Subsidiary objectives were to identify factors affecting the acceptability of, demand for, implementation of, and practicality of the workshop.

# Study 1

## *Method*

### Study design

Workshop 1 was conducted at the University of Bristol on December 8th, 2021. Participants completed questionnaires before and after the workshop, which yielded both qualitative and quantitative data. Questionnaire responses were analysed to understand whether participants found the workshop beneficial, and how to improve it. The study protocol was pre-registered on the Open Science Framework (10.17605/OSF.IO/VN5X9). Ethics approval for the study was obtained from the School of Psychological Science Research Ethics Committee at the University of Bristol (Approval code: 9787). Participants gave informed consent prior to the study using an online consent form.

### Participants

Seven people from the University of Bristol's Faculty of Life Sciences (FLS) attended the workshop (see Table 1 for attendee characteristics). We narrowly missed our pre-specified minimum desirable sample of 8, with our projected ideal target sample being 20-25 attendees. To be eligible, attendees had to be a postgraduate researcher or member of research staff at the University of Bristol, based within one of the five schools within the FLS.

**Table 1.** The seven attendees' career stages, Schools within the University of Bristol, and self-reported research areas.

| Attendee | Career Stage | School | Research Area |
|---|---|---|---|
| Attendee 1 | PhD Year 4 | Physiology, Pharmacology, and Neuroscience | Neuroscience; sleep & neurodegeneration |
| Attendee 2 | PhD Year 3 | Biological Sciences | Machine learning for genomic data |
| Attendee 3 | PhD Year 2 | Psychological Science | Vision |
| Attendee 4 | Research Technician | Cellular and Molecular Medicine | Virology |
| Attendee 5 | PhD Year 3 | Psychological Science | Cognitive neuroscience of language |
| Attendee 6 | PhD Year 5 | Physiology, Pharmacology, and Neuroscience | Sensorimotor systems neuroscience |
| Attendee 7 | PhD Year 4 | Cellular and Molecular Medicine | Regenerative medicine and microbiology |

Participants were recruited through various channels. Administrative support across the five schools within the FLS circulated our recruitment invitation email to postgraduate researchers and staff. Some used specific mailing lists. Others added our recruitment invitation to their weekly bulletin, a newsletter-like email. Some schools did both. These invitations were re-circulated after two weeks. Deviating from our protocol, the invitation was posted on the University of Bristol Tobacco and Alcohol Research Group (TARG) social media channel. The

workshop was also mentioned in the School of Psychological Science's weekly online assembly. Additionally, the invitation was circulated around the University of Bristol's Medical School. We made these deviations to bolster low response rates.

The recruitment email explained what the workshop would entail, and that it was being run as part of our study. Those wishing to attend were asked to contact RC with an expression of interest. Fifteen people contacted RC in total. Five could not make the new date (after the date was changed to avoid coinciding with University and College Union strike action), and three withdrew on the day for personal reasons. No incentives were given for participation.

**Materials**

Study data were collected using two online questionnaires: a pre- and a post-workshop questionnaire. The pre-workshop questionnaire collected predominantly demographic data, such as attendees' research areas and career stages. It also contained a set of five statements, for example: *"I have a good understanding of scientific method"*. Respondents had to rate agreement with these statements on a seven-point scale. The post-workshop questionnaire contained the same set of five statements, plus two additional ones which were used to corroborate whether attendees benefitted from the workshop. Attendees provided their initials on both questionnaires so that their responses across the two could be matched. These were changed to their anonymised identifier by a third party before the analysis. Boxes requesting additional comments were also provided. The rest of the questionnaire asked attendees to rate agreement with statements about specific aspects of the workshop, such as the content and format. Spaces were provided in each section for additional comments. Both sets of questionnaires were developed by the research team and can be found at: https://osf.io/m5vyj/.

The materials used to run the workshop consisted of a set of pre-workshop activities (https://osf.io/aujsp/) and a set of slides (https://osf.io/eug8f/). The pre-workshop activities included recommended reading of a blog post by Paul Spector (Spector, 2021). The central, mandatory component was to watch a video and complete an activity, both of which were created by the research team. The video gave a brief overview of scientific methodology, before providing instructions for the activity. The activity was to create two methodological schematics, pictorially representing what the research process looks like to the individual.

**Procedure**

Before the workshop, RC emailed attendees the pre-workshop questionnaire and activities to complete. The workshop itself was run by RC, who collected field notes throughout the 4-hour session. Upon their arrival, RC photographed the methodological schematics that attendees had brought. RC began by introducing the plan and rationale for the session. Then, attendees spent 30 minutes discussing their methodological schematics in pairs, refining them in the process. RC photographed these updated schematics. After a break, three hours were spent discussing the three types of inference in turn.

An hour was spent covering inductive inference. To start, attendees discussed in pairs what they understand by 'inductive inference' (this was repeated for deduction and abduction). RC then introduced the concept, providing examples and frequently asking questions. Many

examples were elicited and discussed. This ended with an exercise in which attendees were asked to identify inductive inferences in an example used repeatedly throughout the workshop (the link between H. Pylori bacteria and stomach ulcers). Then, attendees considered and discussed how they used induction in their own research, annotating their schematics if they wished. Lastly, attendees discussed how one can be more confident in inductive inferences, both in general terms, and in the specific terms of their own research. This involved considering the problems inherent in induction, as well as the factors which influence how much reliability one wants from inductive inferences.

After another break, seventy minutes were spent covering deductive inference. The deductive section began with a description and examples, with RC asking questions, eliciting examples, and prompting discussion throughout. Then, RC introduced the ideas of hypothetico-deductivism and deductive falsification. Next, attendees identified where deduction was used in the H. Pylori example. This was followed with attendees discussing how they use deduction in their own research, whether they use (or could start to use) deductive falsification in their research, and ultimately, how one can improve deductive inferences. RC ended this section with a quick discussion on where uncertainty arises with deductive falsificationism, and how to use it more effectively.

After the final break, 40 minutes were spent on abductive inference, following the same structure as induction. RC then finished with two brief slides which summarised the three types of inference, highlighting the differences in their argumentative structure, as well as their interconnectedness. This summary narrowly finished within the allotted 4-hours, at which point the workshop had to end. RC apologised that there was not time for a free discussion, photographed the schematics one final time, thanked the attendees, and asked them to complete the feedback questionnaire as soon as possible.

**Analysis**

*Pre-workshop questionnaire.*

RC analysed the two questions *"What education (if any) have you had in Philosophy of Science?"* and *"What motivated you to attend this workshop?"* to help infer to whom the workshop appealed. Using conventional content analysis (Hsieh & Shannon, 2005), RC categorised responses to these questions based on their manifest content.

*Post-workshop questionnaire.*

Attendees rated their agreement with a set of five statements about their understanding of scientific method and inference in both the pre- and post-workshop questionnaires. RC compared responses across these two timepoints to gauge whether attendees understanding had improved. RC used the two additional questions which appeared only in the post-workshop questionnaire to corroborate the validity of these measures, and further infer whether attendees thought that they had benefitted from the workshop.

The remainder of the post-workshop questionnaire contained either text responses, ratings of agreement to statements on a 7-point scale, or a multiple-choice question. RC analysed all text responses using conventional content analysis. He began by grouping all text into codes based

on the research question it helped to answer. Then, he categorised the manifest content within those groupings, drawing primary conclusions from common trends. In cases of unique or minority experiences, RC looked for patterns across those individuals' data, occasionally referring to the field notes to better understand those experiences. After analysing and writing up his conclusions, RC looked back over the entire dataset to check whether he had neglected – or, conversely, over-interpreted – any data. RC's epistemological and interpretive stance for all qualitative analyses is included in the *Supplementary Methods.*

RC analysed the remaining responses on a case-by-case basis, as they all pertained to a specific component of the workshop (e.g., improvements to the content). When possible, he drew conclusions by combining these responses with relevant information from the qualitative content analysis.

*Field notes.*

Lastly, RC used the field notes he made during the workshop as a final source of ideas. These are predominantly drawn on in the discussion when making statements about how to alter the workshop.

## *Results*

We ordered the results to first answer Research Question 1: whether attendees found the workshop beneficial. We then present our data on questions of the acceptability of, demand for, implementation of, and practicality of the workshop. These are all intended to inform Research Question 2: how to maximally improve the workshop for future iterations. In the Introduction of Study 2 we discuss the changes we made to the workshop in light of Study 1.

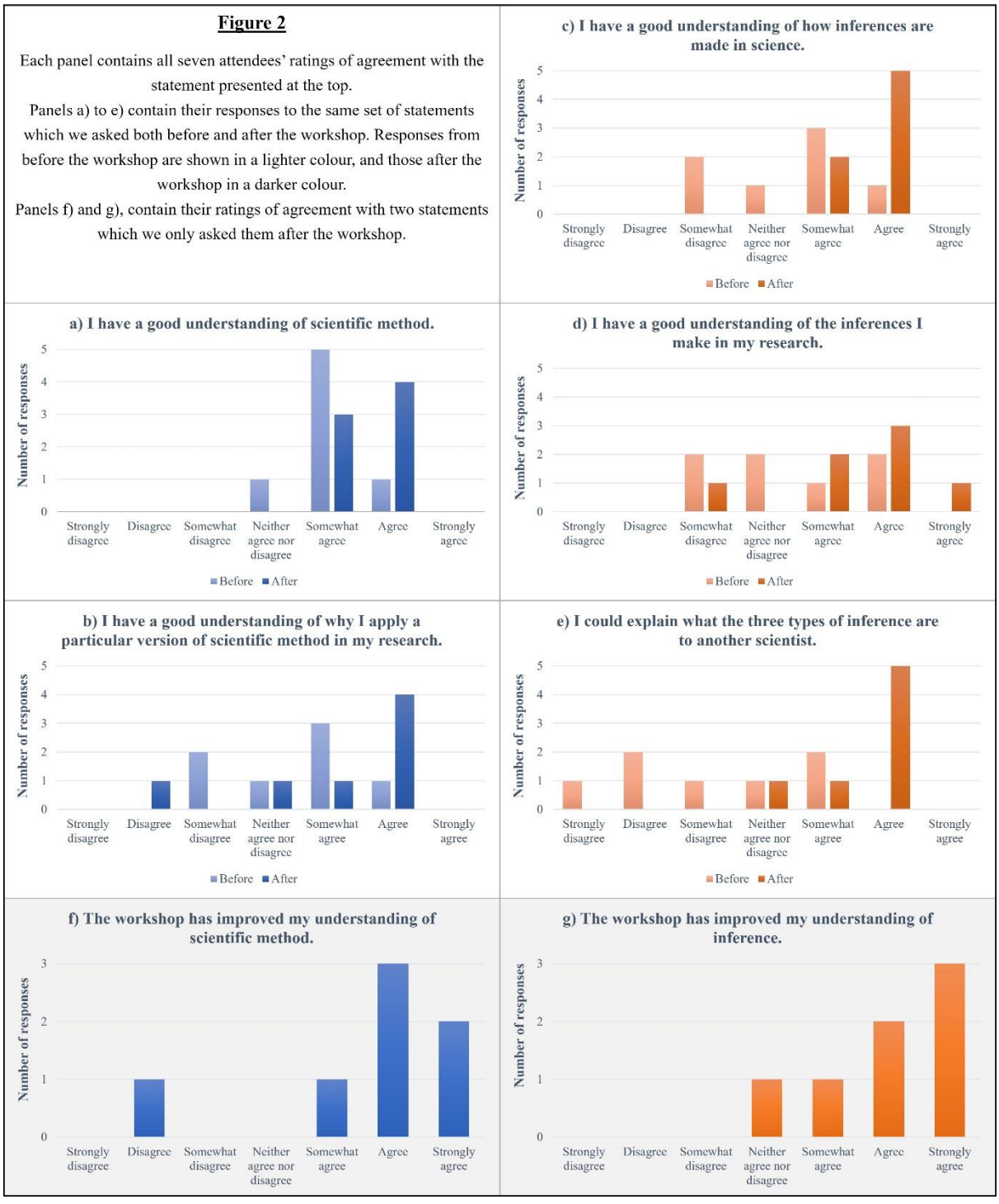
### Benefits of the workshop

Research Question 1 was whether researchers benefitted from attending the workshop. We asked attendees to rate their agreement with a set of five statements before and after the workshop, regarding their understanding of scientific method and inference (Figure 2.a-e). Ratings were given on a 7-point scale, ranging from (1) strongly disagree to (7) strongly agree, with 4 being 'neither agree nor disagree'.

We averaged ratings across participants for the two statements pertaining to scientific method. For these, attendees rated that their understanding had improved by .57 on the 7-point scale. We did the same for the three statements on inference, averaging ratings for all attendees across all three statements. Here, attendees rated that their understanding had improved by 1.57 on the 7-point scale. The most noteworthy improvement was a 2.4-point increase to statement (e): *"I could explain what the three types of inference are to another scientist"*.

These reported increases concurred with ratings to the two additional statements we posed in the post-workshop questionnaire. The average rating to: *"The workshop has improved my understanding of scientific method"* was 5.57 (between 'somewhat agree' and 'agree'. The average rating to: *"The workshop has improved my understanding of inference"* was 'agree' (Figure 2. f and g). Participants' individual ratings to all seven statements are reported in *Supplementary Table S1*.

**Figure 2.** Attendees' self-ratings about their understanding of scientific method and inference, before and after the workshop.




*Attendees' ratings of agreement with seven statements they were asked in the questionnaires before and after Workshop 1. The statements are presented at the top of each panel (a-g). Ratings were given on a 7-point scale, ranging from (1) 'strongly disagree' to (7) 'strongly agree', with (4) being 'neither agree nor disagree'. Statements a-e were asked in both the pre- and post-workshop questionnaires. Statements f and g were only asked in the post-workshop questionnaire.*

To understand any potential benefits in more depth, we asked attendees to write about some additional topics: (i) whether and how their understanding of *scientific method* has improved; (ii) whether and how their understanding of *inference* has improved; (iii) what the *main things* they took from the session are; (iv) whether they would *change any future research practices* in light of the workshop; and (v) whether they would *recommend* it to a friend or colleague.

RC analysed their responses across these five topics and grouped them under three broad themes: (1) *Benefits of methodological schematics*; (2) *Critical thought*; (3) *Expected to learn more.*

***(1) Benefits of methodological schematics***

> *"The workshop confirmed my understanding of scientific method, whereas before I questioned myself a lot."* – Attendee 6

Drawing and scrutinising personal methodological schematics helped attendees better understand scientific methodology, as well as to consider how to improve their own research practices. Attendee 1 noted how it helped to *"contextualise and think about our work"*, while Attendee 3 appreciated the chance "*to learn <u>how</u> we can examine our own research"* (my emphasis). Attendee 5 wrote that it "*provides a good way to organize your thoughts around your own research and be more aware of certain issues that might happen… [and also] to be more critical about it".*

> *"The workshop gave me the opportunity to reflect on the scientific methodology I use to carry on my research. Asking us to draw the two diagrams was definitely a plus: Being aware of each step and their connection to one another is very informative."* – Attendee 2

For Attendee 3, *"the workshop has highlighted how research can start at different points"*, and Attendee 1 had *"never really thought about the many inferences we make at every decision point in the scientific method".* This heightened recognition of *"the many assumptions underlying our work & each inference we draw"* led Attendee 1 to conclude that they *"hope to really take time to examine [their] research process & assumptions and comment upon these".* Likewise, Attendee 2 said that they hoped to *"articulate more in research papers [their] research methodology".*

In a similar vein, the main thing Attendee 6 took from the workshop was "*the need to plan experiments specifically ahead of time, to maximise inference robustness".* The benefits of plotting out the research process ahead of time was echoed by Attendee 3, who "*would definitely recommend the workshop to a colleague starting out in a research project, as [they] think it definitely helps people to get critical and think deeply about conducting research in any discipline".*

***(2) Critical thought***

In their own ways, six attendees discussed how the workshop helped them to be more critical of inferences. This critical appraisal of inferences was achieved by first establishing a *"clearer understanding of what the 3 types of inference are"* (Attendee 1). Indeed, everyone agreed that the workshop improved their understanding of inference. Attendee 3 also noted how it helped

them to recognise the way in which all three types of inference *"are interlinked"*. This baseline understanding was solidified through our *"hands on approach"* (Attendee 2) and in-depth discussions. This facilitated recognition of "*the problems with inferences… [and that] all inferences are fallible"* (Attendee 4).

> *"It helped me in being more critical of my inference methodologies. How to identify them, knowing their limitations and how to mitigate them."* – Attendee 2

> *"I don't think that my research practices will change regarding the experimental procedure, but I think the way in which I assess literature and the questions that arise from it will be judged more critically."* – Attendee 3

In light of the fallibility of inferences, Attendees 1 and 5 both acknowledged the importance of recognising and *"declaring [one's] assumptions"* in hypothesis testing research. By contrast, Attendees 4 and 6 both inferred the importance of triangulating evidence from *"a variety of methods and types of data to make a conclusion"* (Attendee 4). Attendee 6 *"[hoped] to keep in mind the things we discussed in the workshop and begin confirming a finding from various angles with different methods".*

***(3) Expected to learn more***

The final theme relates to one attendee's experience, which differed significantly from the other six, in that it was not consistently positive. In order to interpret these differences of experience, RC drew conclusions from across all of their questionnaire responses. At times, this analysis is explicitly speculative, as well as being informed partly by RC's field notes.

Beyond learning *"the meaning of the three terms abduction, induction, and deduction"*, Attendee 7 had little positive to say about the workshop, concluding with *"I expect more to learn"*. It is hard to draw strong conclusions about why that might be, as they gave quite limited feedback. Their main feedback was that there was *"too much paired conversation"* and that they would have liked *"discussion with [more] specialised people than the audience"*.

> *"All audience are not fully understanding, so I suggest that there are more specialised people to talk with and discuss our gaps in the research than the audience themselves because they also come to the course to learn. So less group discussion would be better."*

These preferences for less group discussion and for discussion with more specialised people were unique. They potentially represent a misunderstanding of what the workshop was about. Rather than being about 'discussing gaps in our research', the workshop presupposed the ability to pictorially map out one's research process before arrival. This might explain why they had not created a methodological schematic before the session.[1]

A large part of the group discussion relied upon the schematics; discussing them at the beginning of the workshop and using them as a tool when we reflected on how we use inference. The post-workshop questionnaire contained two statements about whether it was helpful *"… to discuss [one's] methodological schematic with someone at the session"* and *"…*

[1] *Attendee 7 was able to be identified as the one who did not bring a schematic, because everyone's schematics were photographed and linked to their questionnaire responses*.

*consider ways in which we can improve our inferences".* Attendee 7 'somewhat disagreed' with both statements, while the other 6 attendees either 'agreed' or 'strongly agreed' with them. Again, these differences might be explained by them having not brought a schematic.

Finally, Attendee 7 would have liked to be grouped with people from a similar field. They wrote: *"I think if the workshop is more specific to each field, for example the workshop can be involving more examples to medical field and the audience all grouped as their speciality".* This is helpful, and echoes an early plan of ours to group attendees by discipline. The idea was to first discuss things within disciplines, before fostering cross-disciplinary discussion. Due to the small number of attendees, this did not happen. It is worth pointing out, however, that Attendee 5 benefitted from having to explain their research *"to someone from other areas… who has no idea what [their] research is about".* Consequently, there is evidence in support of both mixed and grouped tables.

**Acceptability**

Overall, it is clear that the workshop was acceptable to everyone who attended, perhaps with the slight exception of Attendee 7. All seven attendees said that they would recommend it to a friend or colleague. Asked whether the workshop content was engaging, four 'strongly agreed', two 'agreed', and one 'somewhat disagreed'. Asked what they did not like, or what could be improved, Attendees 1 to 6 offered constructive feedback about specific aspects (to be discussed) but did not have anything overtly negative to say. On the contrary, they had some very positive things to say. Attendee 3 "*liked meeting the other attendees and sharing our thoughts and research!"*, while Attendee 4 thought that *"Robbie is enthusiastic, and [they] found he made the tricky subjects easier to understand".*

> *"I thought it was a brilliant workshop and even though it was heavy (despite only covering 3 things!) the time flew by. My head was a bit melted by the end, but it was engaging throughout. Robbie is 10/10 teacher, and his content was excellent. I would LOVE for this to be extended to a philosophy of science training course."* – Attendee 1

> *"Robbie did a really good job structuring and explaining the contents. Also, he has the skills to keep his audience engaged. Also, the activities were really useful to understand the inferences and our own research."* – Attendee 5

**Demand**

Despite sharing the workshop invitation across multiple schools, we only received interest from 15 individuals. To better understand the demographics to whom the workshop appealed, RC analysed the seven attendees' responses about their prior education in philosophy of science and what their motivations were for attending.

Six attendees reported having received no education in philosophy of science, three of whom said that they had acquired some knowledge on these topics through *"reading books", "seminars"* and *"podcasts".* Attendee 5 was unique in having covered a *"few bits… in some research methods courses".* Even then, they concluded that their *"approach to philosophy of sciences comes from talks in conferences and discussion with colleagues".*

Motivations to attend were generally split into two, possibly interrelated, categories. Three attendees explicitly wanted to improve their understanding of epistemology, scientific methodology, and philosophy of science (Attendees 1, 4 and 5). Another three focused instead on their desire to reflect on how to improve research (Attendees 2, 3, and 7). The common denominator across these responses is therefore a recognition of the importance and potential benefits of these topics. This may also have been true for Attendee 6. All they expressed, however, was a curiosity *"to look into scientific inference and methodology more closely"*.

**Implementation and Practicality**

*Pre-workshop content: reading*.

In general, attendees were positive about the pre-workshop reading. One attendee 'somewhat disagreed' that it was useful to read before attending, while the others all 'agreed' to some degree that it was useful. Responses were more mixed and less positive, however, about whether it actually helped their understanding of the three types of inference (which was the aim of the exercise). Three attendees expanded on these points, pointing out that they didn't retain much of the information they read.

*Pre-workshop content: video.*

Responses to whether the pre-workshop video helped attendees' understanding of scientific method were mixed, ranging from 'neutral' (2), 'somewhat agree' (1), 'agree' (3), or 'strongly agree' (1). This is fine, as the aim of the video was primarily to contextualise and explain the creation of methodological schematics. Two attendees 'strongly agreed' and 4 'agreed' that it clearly explained what they needed to do for this activity. Nevertheless, despite 'strongly agreeing' with this, Attendee 1 also said that they weren't sure if they were doing it correctly, and *"would have benefitted from having more discussion with the teacher about these things or having them check over the work".* Attendee 7 'somewhat disagreed' that the video clearly explained the activity, which might explain why they did not bring a schematic.

*Pre-workshop content: methodological schematic.*

The methodological schematics achieved their purpose. Attendees 'agreed' that the schematics helped them to visualise their research process, and that it was helpful to draw both abstract and concrete schematics. Likewise, attendees 'agreed' that it was helpful to discuss them with someone at the session. Attendee 2 put it succinctly: *"Very useful- watching the video is not enough. Drawing your own research is testing and honing your understanding".*

*Workshop content.*

All attendees 'agreed' that the workshop content was clear. With the exception of Attendee 7, everyone either 'agreed' or 'strongly agreed' that it was helpful to consider how they use inference in their research. These thoughts were already echoed in much greater depth in the *'Critical thought'* theme.

Feedback on our recurring H. Pylori example was similarly positive. Two attendees offered useful feedback on how the example could be improved:

> *"I think the H. Pylori example was good, but the vagueness of some of the statements perhaps led us down the wrong tracks when discussing… Perhaps more detail in this example would be helpful? But it was still really good."* – Attendee 3

> *"On the concrete schematic that was constantly used as an example, observations and beliefs were collapsed. Following the logic of induction (and even abduction), those two might be perfectly separated (like a weaker version of data and theory)."* – Attendee 5

We did not have time for a free discussion, so it is hard to comment on. Attendee 5 was "*pretty sure it would help the workshop"*, and we are inclined to agree.

*Workshop format.*

Overall, the discursive format – involving lots of paired work – appeared largely successful. People 'agreed' that their partners were engaged, and that the group sizes were appropriate to the tasks. Attendee 6 also *"liked the small group size, as it allowed everyone to contribute".* We hoped to run the workshop in person because of this discursive element, and so it is reassuring that there was a 6:1 preference for having the workshop in-person as opposed to online.

While the pace of the session was rated 'about right', the workshop's length was an issue. Indeed, the need to alter the length of the workshop was the biggest lesson on format. Five attendees found it 'a bit too long', 1 'too long', and only 1 'just right'. Overall, attendees appreciated that the *"session was an appropriate length"* (Attendee 3) relative to the amount of content. Attendee 6 wrote: *"There's a lot to discuss and I can't think of anything I'd remove to make it shorter".* Given that we did not have time for a final discussion, four hours is clearly insufficient.

Five attendees 'agreed' to some extent that they would have preferred it to be split into two sessions. Thoughts on logistics of this differed, however. Attendee 4 felt that "*it would work better as two slightly shorter sessions"* (my emphasis), whereas Attendee 5 thought that if it was split into two sessions *"each session might need to last 3 hours… because students might need some reminders of the previous session".* Attendee 6 had a slightly different concern, *"[worrying] about losing that engagement if we went away and came back another day".* We discuss how we incorporated this feedback in the next section.

## Study 2

### *Introduction*

After Workshop 1 we made some changes to the workshop's content and format. These were intended to make the large amount of content more manageable, while also making the applied approach more accessible. The only major change involved splitting the workshop into two, three-hour sessions. We decided to cover scientific methodology and create methodological schematics in session 1, rather than before the workshop. The aim was to remove the possibility of people arriving without a schematic. Moreover, we thought it would be beneficial to have the facilitator and other attendees at hand to offer support where needed, following feedback from Attendee 1.

Splitting the workshop allowed us to break up the content on inference across the two sessions. In addition to stopping people's brains from becoming *"a bit melted"* (Attendee 1), we wanted to have ample time for free discussions at the end of each session. Furthermore, we felt that a week's break would give attendees time to reflect on what they had learned in session 1. Better still, we hoped that revisiting session 1's content in session 2 would help to consolidate understanding and improve overall information retention. These are all of the major changes made after Workshop 1. Complete documentation of every change can be found in our Protocol Amendment: https://osf.io/qdgnx/.

## *Method*

### Study design

Workshop 2 was conducted at the University of Bristol across two sessions on March 23rd and 30th, 2022. Participants completed questionnaires before and after the workshop, almost the same as in Study 1, except for a few minor modifications. The study protocol was updated for Study 2 and pre-registered on the Open Science Framework (10.17605/OSF.IO/VECB9). Ethics approval for the study was obtained from the School of Psychological Science Research Ethics Committee at the University of Bristol (Approval code: 10419). Participants gave informed consent to prior to the study using an online consent form.

### Participants

Five people from the University of Bristol attended the workshop (see Table 2 for attendee characteristics). This narrowly missed our desired sample of 6-16 people. To be eligible, attendees had to be a postgraduate researcher or member of research staff at the University of Bristol, based either within the FLS or the Centre for Doctoral Training (CDT) in Digital Health and Care.

**Table 2.** The five attendees' career stages, locations within the University of Bristol, and self-reported research areas.

| Attendee | Career Stage | School or Faculty | Research Area |
|---|---|---|---|
| Attendee 8 | PhD Year 2 | Faculty of Life Sciences | Marine bioacoustics |
| Attendee 9 | PhD Year 3 | School of Engineering (Digital Health) | Psychiatric epidemiology and machine learning |
| Attendee 10 | Master's by Research | School of Psychological Science | Low-level vision |
| Attendee 11 | PhD Year 4 | School of Psychological Science | Psychology and nutrition |
| Attendee 12 | PhD Year 1 | Faculty of Life Sciences | Forest ecology |

Our recruitment methods were similar to Study 1. Emails were circulated around the FLS and CDT in Digital Health and care twice, two weeks apart. Adverts were put in the weekly bulletins in each school within the FLS. Two emails were sent to the postgraduate-specific mailing list for the FLS. Five individuals, who had expressed interest in Workshop 1 but could not attend, were re-contacted. The recruitment email was modified to reflect the changes to the

workshop format and monetary reimbursement. Twelve people contacted RC in total. Three could not make the dates, and four cancelled near the day due to personal reasons. Reimbursement of £30 was given to each attendee.

**Materials**

The materials changed only slightly after Workshop 1. The full list of changes can be found in our *Supplementary Methods*, and the updated questionnaires and slides can be found: https://osf.io/knqpm/ and https://osf.io/ef76g/. Three changes to outcome measures are worth mentioning. We changed Statements (b), (f), and the follow-up to Statement (f) in the questionnaires. In Workshop 1, these statements had pertained to understanding of *scientific method*. In Workshop 2, we changed them to be about understanding of *one's research process*.

**Procedure**

The procedure remained very similar to Workshop 1. Now split into two sessions, content from the pre-workshop activities was included in session 1. Session 1 ran as follows. RC introduced attendees to the workshop and gave a brief presentation on scientific method. Then, attendees had 30 minutes to draw their schematics before spending another 30 minutes discussing them in pairs. After a break, one hour was spent on inductive inference. After a second break, 20 minutes was spent on a free discussion, during which time attendees discussed remaining thoughts or questions.

Session 2 started with a quick recap of inductive inference. Then, one hour was spent discussing deductive inference, followed by a break. Next, abductive inference was covered for 45 minutes, followed by another break. The final free discussion and wrap up took around 20 minutes, allowing us to finish 25 minutes ahead of schedule.

**Analysis**

Analysis of the pre-workshop questionnaire, post-workshop questionnaire, and field notes taken by RC during the workshop were the same as in *Study 1*.

## *Results*

As with Study 1, we ordered the results to first answer Research Question 1: whether attendees found the workshop beneficial. We then present data on the acceptability of, demand for, implementation of, and practicality of the workshop. This helps to inform Research Question 2: how to maximally improve the workshop.

**Benefits of the workshop**

To inform Research Question 1, we presented attendees the same set of statements as in Study 1 (except statement b) before and after the workshop (Figure 3. a-e). Attendees rated agreement with the statements on a 7-point scale, ranging from (1) strongly disagree to (7) strongly agree. Mirroring the positive trend observed in Workshop 1, attendees' ratings improved on average by around 1 point on the scale for statements (a) to (d). Again, the biggest shift was for statement (e) *"I could explain what the three types of inference are to another scientist".* Here,

the mean rating went from 3.6 before the workshop—halfway between ‘somewhat disagree’ and ‘neither agree nor disagree’— to 6.8 after, with 4 of 5 attendees ‘strongly agreeing’.

As in Study 1, these trends concurred with the two additional statements we posed in the post-workshop questionnaire. Statement (f) was *“The workshop has helped me to gain a clearer picture of my overarching research process”*. Statement (g) was *“The workshop has improved my understanding of inference”*. To both statements, 4 of 5 attendees ‘strongly agreed’, while one ‘agreed’ (Figure 3. f and g). Participants’ individual ratings to all seven statements are reported in *Supplementary Table S2*.

**Figure 3.** Attendees' self-ratings about their understanding of scientific method, their research processes, and inference, before and after the workshop.


*Attendees' ratings of agreement with seven statements they were asked in the questionnaires before and after Workshop 2. The statements are presented at the top of each panel (a-g). Ratings were given on a 7-point scale, ranging from (1) 'strongly disagree' to (7) 'strongly agree', with (4) being 'neither agree nor disagree'. Statements a-e were asked in both the pre- and post-workshop questionnaires. Statements f and g were only asked in the post-workshop questionnaire.*

We also asked attendees to write about the same five topics as in Study 1 (i – v), except for (i). Rather than asking whether and how the workshop had improved their understanding of *scientific* method, we asked whether and how the workshop had helped them to gain *a clearer picture of their overarching research process*.

RC found that responses to these topics could be organised into the same two themes as in Study 1: (1) *Benefits of methodological schematics*, and (2) *Critical thought.*

***(1) Benefits of methodological schematics***

All five attendees 'strongly agreed' that drawing a schematic helped them to visualise their research processes. They also agreed that it was helpful to discuss them with someone else at the session. Attendee 11 felt that this discussion would help them in *"communicating [their] research to the wider community".*

> *"The materials from the workshop helped me develop a greater understanding of my research process, of my thesis. It helped me create a clearer and easier to explain overview of the key theories, observations, and key findings from my PhD."* – Attendee 11

Drawing the schematic helped Attendee 10 *"clarify the different stages of the scientific process, and the interactions between each stage"* (Attendee 10). These benefits were echoed by Attendee 9, who found it *"helpful to map out [their] whole research question and split it into separate parts".* Breaking the research process into these discrete steps also helped attendees to recognise the central importance of literature at the beginning of the cycle. This prompted some of the critical thought which we discuss in the next section.

***(2) Critical thought***

As in Workshop 1, we met our baseline aim of ensuring that everybody gained the vocabulary and understanding of the three types of inference, with all five attendees writing about how the workshop helped in this regard. This necessary step provided the foundations for us to achieve our main goal of facilitating critical thought. We achieved this aim in several regards, much of which was centred on the research design phase. For example, three attendees noted how wider literature plays an integral role at the beginning of the research process, often determining what future research is conducted. Recognising this, they wrote about how they will be *"more critical"* of literature in the future. This manifested in assessing *"the quality of the literature supporting your inferences"*, *"conflicting evidence"* (Attendee 8), and also the methodologies employed:

> *"I think there are a lot of 'I did this because a past paper found X / X is widely accepted' type research practices, and that approach may benefit from some scrutiny!"* – Attendee 9

In addition to the assumptions inherent in the literature, four attendees also recognised that *"we have many implicit and explicit assumptions in our work"* (Attendee 10), which are important to acknowledge. Attendee 11 reported aiming to *"acknowledge [their] assumptions and the assumptions imbedded within the literature cited, and provide clearer explanations of [their] research processes, including when [they] would reject or accept hypothesis".*

Two attendees also discussed how they would be more critical of their question formation in the future. Attendee 9 stated that *"the workshop has made [them] reflect on hypothesis generation and the steps that feed into it, and how important soundness of premises are".* Similarly, Attendee 8 reported how the workshop *"allowed [them] to think about what [they are] actually inferring or deducing from the theory/literature base to form [their] questions".* Relatedly Attendee 8 also mentioned how *"in designing [their] own questions, [they] will give more thought to why it's a good question to ask, and how [they] can collect and analyse data in a transparent way that has the maximum value possible".*

Attendees also discussed the value of critical thought at the analysis and reporting stages of research. Three mentioned how the workshop discussion gave rise to encouraging *transparency*. Attendee 11 mentioned being *"more transparent/reflexive on how [they] make decisions about [their] research or inferences".* Contextualising one's research within the wider literature was the other main trend. For example, Attendee 8 said that the workshop *"has helped [them] think about the context and descriptive way [they] will place the research in the current literature once complete".*

**Acceptability**

Feedback on the acceptability of the workshop was very encouraging. All attendees would recommend the workshop to others, and 'strongly agreed' that the content was engaging. Ratings towards the merit of the workshop's content and format were high, and any critical feedback pertained constructively to minor improvements.

> *"Physically writing down the research process and visualising it was really helpful. It was also hosted in a friendly and non-intimidating way, which made a big difference to someone a little more introverted/shy. I liked that the other attendees were from different disciplines, it gave me more perspective to how inference can occur more widely than just in my own experience. I also liked the cats."* – Attendee 8

> *"I enjoyed hearing about everyone else's research and meeting other people, especially after all the lockdowns! I also really enjoyed the informal feel of the session, it felt like a space where you could just think aloud and bounce ideas off each other without feeling self-conscious."* – Attendee 9

> *"Robbie's style was very good, explaining in layman's terms at the start but then dropping into the specific language of philosophy as required."* – Attendee 10

**Demand**

As in Study 1, demand for the workshop was poor. Despite wide recruitment, there were only 12 expressions of interest. RC analysed the attendees' reported educations in philosophy of science, and their motivations to attend, to better understand who the workshop appealed to.

Attendees reported some previous acquaintance with philosophy of science, but nothing particularly in-depth. Attendee 8 had *"completed an A-level in Philosophy of Science"*, while Attendee 10 had done *"limited reading in the general area".* Attendees 11 and 12 claimed to

have had no education in philosophy of science. Attendee 9 had the most, but nevertheless, only an introductory module.

> *"I think I did an introduction to Philosophy of Science module at undergrad - I can't remember much except perhaps logical inference."* – Attendee 9

Four attendees mentioned a desire to learn about the broad topic of scientific methodology. Attendee 8 was interested to see *"how a 'big picture' understanding aids your own ability to do very specialised research".* Attendee 9 linked this desire to *"zoom out"* to their interest in *"interrogating your scientific inferences"* and applying *"this kind of critical thinking"* to their area of interest. Two attendees mentioned a desire to learn about inference specifically, with Attendee 10 wanting *"to properly understand the difference between induction and deduction and related ideas"*.

**Implementation and Practicality**

*Workshop content.*

Overall, feedback about the content was extremely positive. Everyone unanimously agreed that the workshop's fundamental components were helpful (drawing and discussing the schematics; the descriptions and examples; considering one's inferences and how they could be improved; and the free discussions). Three attendees mentioned specifically how the examples were simple and clear, contributing to their understanding of that type of inference. Feedback about the recurring H. Pylori example was generally positive, although two attendees mentioned that it could do with improvement. Attendee 9 felt that the jump from simpler explanations to that example was quite big, suggesting another example to potentially bridge the gap. Similarly, Attendee 10 felt that the example *"was overall very helpful, but could probably do with a bit of refinement to make sure the terminology exactly ties up with the teaching"*.

The other suggestions for improvement pertained to linking the content directly back to research practice. Attendee 8 thought that *"it might have been helpful to have examples of how people have directly applied the concepts of the workshop to a PhD project or study".* Likewise, Attendee 9 said that they would find it helpful *"knowing how the things we learned might impact real life research practice… Perhaps an invitation in the discussion to be more critical of how current research practices may be bypassing these steps or making truly questionable inferences".*

*Workshop format.*

The discursive format, with frequent engagement in paired discussion, received positive feedback from everyone. Attendee 12 *"liked the idea of sharing within smaller groups and relating to [their] own research work".* Attendees enjoyed *"hearing about [their] partner's research"* (Attendee 9), and moreover felt that the discussions *"improved the communication of [their] own research… [with] the chance to get feedback as well"* (Attendee 11). All five voted for a preference towards 'in person' rather than 'online'. Attendee 11 tied this to the benefits of *"interaction and discussion with others [which] really contributed to [their] understanding of the materials and to [their] experience of the workshop".*

Study 1 taught us that the workshop was probably better with small groups, giving more people the chance to speak, while also creating a more intimate feel. These conclusions were reinforced by Attendee 8, who appreciated the *"personal approach"* and who felt that *"a much larger group would have stopped [them] from contributing as much".*

Others echoed their appreciation for the feel of the session. Attendee 11 noted how the *"session felt very comfortable, [and that] it was easy to talk to people".* Nevertheless, five attendees was perhaps a bit too small. Attendee 9 felt that *"an extra couple of people and the group size would have been perfect, but we definitely had enough people to have good discussions".* This at least suggests that, even with only five people, the workshop can function successfully in the discursive manner intended.

The major format change of splitting the workshop into two, smaller sessions was a success. Both sessions finished with time for free discussions, plus extra to spare. Other than 1 vote of 'a bit too fast' for the pace of session 2, everyone felt that the lengths and paces of both sessions were 'just right'. This suggests that we got the right balance of content into each session. We also asked whether attendees would have preferred the sessions to have been split across the morning and afternoon of the same day. There was general consensus that this was not preferable; one attendee was neutral, the other four were on the side of 'disagree'. Given the chance to elaborate, two attendees noted the benefits of a week's reflection.

> *"The time in between the sessions to go back to my work and think about inference for a week before the second session gave me more perspective than if we had considered it all in one day"* – Attendee 8

There were four, minor suggestions of improvement. Two people mentioned how it would have been nice to have had coffee. Attendee 10 would have liked *"more examples [on each type of inference] to work through, to be absolutely sure [they] had the logic nailed".* Attendee 9 *"may have needed more time to really think about what type of inference each step was making, but [still] managed to identify the different types of inference at at least a few stages in [their] schematic".* And finally, Attendee 8 *"found it initially a bit difficult to understand and give feedback on [their] partner's schematic in the time available, because it was a discipline really unfamiliar to [them]… [they] would have been more able to provide feedback if [their partner] had focussed on a single chapter/study".* We will endeavour to incorporate all suggested changes in the future.

## Discussion

In summary, both workshops appeared to benefit each attendee in varying ways and to different degrees. Diagrammatically breaking down the research process into discrete steps and seeing how they are all interlinked helped almost every attendee to visualise and scrutinise how they carry out their research. The act of creating such a schematic appeared to be novel for all attendees. A few suggested that they might use this activity in future. Furthermore, the systematic approach to learning about inference gave all attendees greater confidence in their understanding of these concepts. This baseline understanding helped attendees to recognise and mitigate some of the problems inherent in making inferences, both on an abstract level, and with the inferences they make in their own research. These conclusions are very

encouraging, they suggest that the workshop was beneficial in the way we had hoped. Moreover, they have helped us to identify other potential outcomes. These are outlined in Figure 4, below.

**Figure 4.** Logic model for the workshop, showing the pathway from the workshop's core components to the key outcomes we hope to achieve.

| Input | Activity | Output | Outcome |
|---|---|---|---|
| *The necessary components to run the workshop are:* | *The core workshop activities are:* | *The immediate aims of the activities are to:* | *Long term we hope that these activities will promote:* |
| * **a facilitator** who understands the content and delivery style.<br><br>* 4+ **researchers** with research projects that they can pictorially represent.<br><br>* necessary **materials** (presentation slides, A3 paper, coloured pens).<br><br>* necessary **facilities** (a room with tables for paired work and a large monitor for the slides). | * a brief presentation on scientific methodology.<br><br>* pictorial representation of individual research processes (attendees create methodological schematics).<br><br>* three discursive seminars on inference (induction, deduction, and abduction – respectively).<br><br>* free discussions. | * convey ideas about scientific methodology and three types of inference.<br><br>* help attendees to visualise their research processes and the inferences made therein.<br><br>* teach the rationale and specific problems of each inference methodology.<br><br>* promote critical thought about research, inference, and epistemology. | * understanding of the logic behind inferential fallibility and the need for epistemic humility.<br><br>* appreciation of the many assumptions in research; the need to recognise and minimise them.<br><br>* appreciation of such things as: triangulation; methodological transparency; replication; publishing null results.<br><br>* corresponding changes in research behaviours. |

*This model was developed after analysing both studies, combining our preexisting ideas with lessons from the workshops.*

Our discursive and applied approach to learning appeared to be effective for 11 out of 12 attendees. We employed this 'hands-on' method for several reasons. Fist, meta-analytical evidence suggests that STEM students learn more when courses employ an active as opposed to 'lecture-centric' approach. (Freeman et al., 2014). Second, we wanted attendees to discover for themselves how these abstract epistemological concepts are important for their own work. This thought coheres with past research on teaching epistemology to physics students (German grades 11-13). Meyling (1997) argued that epistemology should be integrated *into* science education, learned about through practical application, as it applies to a given topic. Our workshop invited attendees to break down their research processes and interrogate the steps through which they generate knowledge. By doing this in collaboration with researchers from other disciplines, we aimed to elucidate how disciplinary differences shape the epistemological and methodological practices of those disciplines. These practices are recommended by others who discuss the benefits of engaging interdisciplinary researchers in epistemological and 'meta-cognitive' reflection (Boon & Van Baalen, 2018). Given the predominantly positive feedback and results from this workshop, our theoretical perspective of employing active learning remains unchanged.

*Methodological considerations.*

Although our data provide encouraging evidence for the workshop's benefits, there are some important caveats. First, our attendees represented a self-selected and highly motivated sample.

Most attended the workshop with preexisting beliefs that they would benefit from improving their understanding of scientific methodology or inference. To some degree, attendees' positive responses may reflect their interest in the subjects and immediate enthusiasm from participating in the workshop. Moreover, attendees established cordial relationships with RC during the workshops, which were intimate due to the small group sizes. The feedback questionnaire asked questions with clear demand characteristics regarding what the researcher would hope to hear (e.g., that attendees found the content 'engaging'). Additionally, attendees likely suspected a chance of identification from their questionnaire responses, despite anonymisation, due to the small number of respondents. Combined, these factors should make us cautious of over-interpreting the positivity of the workshop's benefits.

It is important to determine whether the ideas and enthusiasm reported within two days were retained longer-term. Specifically, we do not know whether the workshop influenced any future research or analytical practices, and if so in what way. Attendees reported ways they believed or hoped the workshop might alter their thoughts on conducting or conceptualising future research. Learning whether these thoughts came to fruition would much more highly attest to the utility of the workshop. In a future study, we will collect data on distal measures such as changes in understanding and behaviours.

Prior to Study 1, we pre-specified a minimum desirable sample of 8, and ideal sample of around 20, because we wanted to investigate whether attendees benefitted from the workshop differently, depending on their research areas. Unquestionably, a larger sample would have conferred a greater volume and diversity of experiences to analyse. But more importantly, it would have afforded the nuance to assess whether the workshop appeals to and affects researchers from different disciplines in various ways. This information would be helpful to gain in the future. For example, we may want to tailor the workshop to different disciplines, utilising more discipline-specific examples, as Attendee 7 suggested. Moreover, larger groups could enable us to group people by their research disciplines. This could be very useful given the discursive nature of the workshop. In-depth, critical discussions about one's research may be easier and more fruitful when paired someone with a similar methodological specialisation. Ultimately, running the workshop with only 7 and 5 people respectively taught us the value of having fewer attendees than planned. Even in a spacious room, having more than around 14 people is probably not desirable because one risks making the workshop noisy and more impersonal.

*Future directions of the workshop.*

A key lesson from both studies is that it is vital to give attendees ample time to think and discuss things in pairs. Given that there was spare time at the end of each session in Study 2, this should be perfectly feasible. When learning the fundamentals of each inference type, attendees need to be allowed to consider lots of examples among themselves. This should help them to gain confidence in their understanding in an unintimidating way. The same could be said for the free discussion at the end. For the free final discussions, it may be worth asking attendees to discuss in pairs their 'main take-homes', or 'potential changes to future practice', before eliciting their thoughts in front of the wider group. As Attendee 9 suggested, this could include

an invitation to critical discussion on how people sometimes bypass the steps in the research process, and make questionable inferences.

Recruiting sufficient numbers was a concern across both studies. To improve future participation, we need to convey the workshop's potential benefits more clearly. We compared attendees' motivations to attend across Studies 1 and 2. The emphasis they placed on why they wanted to attend reflected the differences in the recruitment emails. In Study 1, the recruitment advertisement emphasised 'understanding scientific method and inference'. Correspondingly, motivations to attend were grounded in a desire to learn more about these topics. By contrast, in Study 2 the advertisement emphasised that *"you will visually map out your research processes… [step back and] conceptualise the whole research process"*. Motivations to attend Study 2 focused more on a desire to 'zoom out', with only two attendees mentioning inference specifically. To appeal to a wider audience, future advertisements should incorporate more detail on the benefits experienced by attendees in these studies, reflecting the varied ways that the workshop might help different researchers.

Increasing recruitment numbers will be crucial, as our next step will be to measure the benefits of the workshop using a wider range of measures. We will collect both proximal and distal measures on attendees' attitudes, knowledge, and behaviours. Further, we will collect different types of measures than self-report data, to mitigate issues around interpreting self-report data. This will provide a stronger base of evidence to justify the adoption of the workshop more widely. Moreover, we hope that further testing will shed light on further improvements to the workshop, or even generate ideas for entirely new resources. All of this is working towards the ultimate aim of developing a set of openly accessible workshop materials which others can freely adapt and use. To this end, we have developed an implementation plan for those who might wish to run this workshop themselves (see *Supplementary Information*).

## Conclusion

Both studies provided encouraging evidence that the workshop can improve researchers' understanding of inference, and help them to visualise their research processes better. The combined activities appear to stimulate scepticism and critical thought across the research process, particularly at the study design phase. Confidence in these conclusions should be tempered in light of the low sample sizes and corresponding possibilities that the self-report data may be biased towards positive responses. Nevertheless, RC's experience as facilitator is that, at the very least, attendees enjoyed themselves and left having had stimulating discussions. Whether the immediate enthusiasm towards the ideas that were discussed persisted into the future is something we hope to test in a future trial, along with other distal measures. This will help us to assess the potential benefits of the workshop more rigorously.

## Data availability

All study data and materials are hosted on the University of Bristol's online data repository (data.bris) as open data at: **https://doi.org/10.5523/bris.15946fz1h1f862m5gz2p5jo000.**

Additional resources – including the Supplementary Information and future project updates – are available on our Open Science Framework page: https://doi.org/10.17605/OSF.IO/Y9ZFH

## Contributions and Authorship

Contributor roles are stated in the table below. For items where contributors had similar contributions names are listed in reverse alphabetical order (Z-A).

| | |
|---|---|
| Conceptualisation: | lead Robbie Clark; supporting Marcus Munafò, James Ladyman[2] |
| Data curation: | Robbie Clark |
| Formal analysis: | Robbie Clark |
| Funding acquisition: | Robbie Clark |
| Investigation: | Robbie Clark |
| Methodology: | lead Robbie Clark; supporting Marcus Munafò, James Ladyman |
| Project administration: | Robbie Clark |
| Resources: | lead Robbie Clark; supporting Marcus Munafò, James Ladyman |
| Software: | *Not applicable* |
| Supervision: | Marcus Munafò, James Ladyman |
| Validation: | *Not applicable* |
| Visualisation: | Robbie Clark |
| Writing – original draft | Robbie Clark |
| Writing – review and editing | Marcus Munafò, James Ladyman, Robbie Clark |

## Funding

This project is funded by a Southwest Doctoral Training Partnership (SWDTP) Economic and Social Research Council (ESRC) PhD studentship, which provides RC's stipend. Marcus Munafò is supported by the MRC Integrative Epidemiology Unit at the University of Bristol (MC_UU_0011/7).

## Competing interests

As designers of the study and materials, all three authors have a vested interest in the workshop's desirability and success.

## Acknowledgements

Many thanks to all twelve attendees for their interest and engagement in the workshops. The depth of concentration and discussion was wonderful to see. Without their detailed and considered feedback, this manuscript would not have been possible. A big thank you also to RC's friends in the PhD office, who helped him to pilot an earlier version of the workshop. Their feedback greatly shaped the workshop materials, while also bolstering RC's confidence that the workshop would benefit the attendees.

## References

Albert, K., Brundage, J. S., Sweet, P., & Vandenberghe, F. (2020). Towards a critical realist epistemology? *Journal for the Theory of Social Behaviour*, *50*(3), 357–372. https://doi.org/10.1111/jtsb.12248

Boon, M., & Van Baalen, S. (2018). Epistemology for interdisciplinary research – shifting philosophical paradigms of science. *European Journal for Philosophy of Science*, *9*(1), 16. https://doi.org/10.1007/s13194-018-0242-4

Derksen, M. (2019). Putting Popper to work. *Theory & Psychology*, *29*(4), 449–465. https://doi.org/10.1177/0959354319838343

Eigenbrode, S. D., O'rourke, M., Wulfhorst, J. D., Althoff, D. M., Goldberg, C. S., Merrill, K., Morse, W., Nielsen-Pincus, M., Stephens, J., Winowiecki, L., & Bosque-Pérez, N. A. (2007). Employing Philosophical Dialogue in Collaborative Science. *BioScience*, *57*(1), 55–64. https://doi.org/10.1641/B570109

Fjelland, R. (2022). Teaching Philosophy of Science to Science Students: An Alternative Approach. *Studies in Philosophy and Education*, *41*(2), 243–258. https://doi.org/10.1007/s11217-021-09802-8

Flis, I. (2019). Psychologists psychologizing scientific psychology: An epistemological reading of the replication crisis. *Theory & Psychology*, *29*(2), 158–181. https://doi.org/10.1177/0959354319835322

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Grüne-Yanoff, T. (2014). Teaching philosophy of science to scientists: Why, what and how. *European Journal for Philosophy of Science*, *4*(1), 115–134. https://doi.org/10.1007/s13194-013-0078-x

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Meehl, P. E. (1993). Philosophy of Science: Help or Hindrance? *Psychological Reports*, *72*(3), 707–733. https://doi.org/10.2466/pr0.1993.72.3.707

Meyling, H. (1997). How to Change Students‘ Conceptions of the Epistemology of Science. *Science & Education*, *6*(4), 397–416. https://doi.org/10.1023/A:1017908916810

Morawski, J. (2019). The replication crisis: How might philosophy and theory of psychology be of use? *Journal of Theoretical and Philosophical Psychology*, *39*(4), 218–238. https://doi.org/10.1037/teo0000129

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Almenberg, A. D., Fidler, F., Hilgard, J., Kline, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L., Schönbrodt, F., & Vazire, S. (2021). *Replicability, Robustness, and Reproducibility in Psychological Science*. PsyArXiv. https://doi.org/10.31234/osf.io/ksfvq

O’Rourke, M., & Crowley, S. J. (2013). Philosophical intervention and cross-disciplinary science: The story of the Toolbox Project. *Synthese*, *190*(11), 1937–1954. https://doi.org/10.1007/s11229-012-0175-y

Reiter, B. (2013). *The Epistemology and Methodology of Exploratory Social Science Research: Crossing Popper with Marcuse*. 18.