Methodological triangulation in empirical philosophy (of mathematics)

Löwe, B.; Van Kerkhove, B.

Link to publication

Methodological triangulation
in empirical philosophy (of mathematics)

Benedikt Löwe[1,2,3] & Bart Van Kerkhove[4,5]

March 26, 2018

[1] Institute for Logic, Language and Computation, Universiteit van Amsterdam, Postbus 94242, 1090 GE Amsterdam, The Netherlands; `b.loewe@uva.nl`.

[2] Fachbereich Mathematik, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany.

[3] Churchill College & Faculty of Mathematics, University of Cambridge, Storey's Way, Cambridge CB3 0DS, United Kingdom.

[4] Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium; `bvkerkho@vub.ac.be`.

[5] Vakgroep Gedragswetenschappen, Communicatie en Linguïstiek, Universiteit Hasselt, Campus Diepenbeek, Agoralaan, Gebouw D, 3590 Diepenbeek, Belgium.

Abstract

We reflect on the use of empirical methods in philosophy and the variety of different empirical methods, emphasising that the experimental method in the strict sense is only one of them. Based on our discussion, we argue for the use of methodological triangulation in empirical philosophy.

1 Introduction

This paper is a reflection on using empirical methods in philosophy. Most of the issues discussed in this paper are general and not restricted to any particular branch of philosophy; however, since both authors are philosophers of mathematics, our point of departure and guiding motivation will be the philosophy of mathematics, in particular the philosophy of mathematical practice. We shall observe that this research area has a number of methodological issues in common with *Experimental Philosophy*, explore these, and make a case for a *multiple-method approach*, also known as *methodological triangulation*.

In §§ 2 & 3, we give a brief overview of the two research areas *Philosophy of Mathematical Practice* and *Experimental Philosophy*, respectively. We shall observe that the name "experimental philosophy" suggests that the latter exclusively uses the experimental method rather than empirical methods in general; this observation leads us to a reflection on the term "experimental", issues of nomenclature, and the plurality of methods in empirical research in § 4. We highlight (in § 5) that the plurality of methods used in research is particularly important for empirical research in fields where the isolation and control of variables is either difficult or undesirable; in the social sciences, this paradigm is known, following Campbell and Fiske (1959), as *methodological triangulation*, and

---

we discuss it in § 6. Finally, in § 7, we return to the philosophy of mathematical practice and give examples of the use of triangulation in our field.

2   Philosophy of mathematical practice

Philosophy of Mathematical Practice is an approach towards philosophy of mathematics that takes the practice of mathematicians into account. Its research community is motivated by the fact that the foundational debates in the philosophy of mathematics in the early 20th century had resulted in a focus on a highly idealised version of mathematics, often ignoring the practices of working mathematicians.[1] The titles of volumes such as *What is mathematics, really?* (Hersh, 1997) or *Towards a philosophy of real mathematics* (Corfield, 2003) express the frustration of the respective authors with a philosophy of mathematics dealing with a strongly sanitised version of the discipline. The conferences on *Perspectives on Mathematical Practices* in Brussels, organised by Jean Paul Van Bendegem and the second author, provided a venue for equally aggravated philosophers to meet with researchers from various empirical disciplines studying mathematical practice (cf. Van Kerkhove and Van Bendegem, 2007; Van Kerkhove, 2008; Van Kerkhove et al., 2010); Mancosu (2008) used the term *The Philosophy of Mathematical Practice* as the title for his edited volume that brings together authors "joined by the shared belief that attention to mathematical practice is a necessary condition for a renewal of the philosophy of mathematics" (Mancosu, 2008, p. 2). In 2009, the research community formed the *Association for the Philosophy of Mathematical Practice* (APMP; with Mancosu and Van Bendegem among the nine founding members) which states its aims and scope as follows:

> Over the last few years approaches to the philosophy of mathematics that focus on mathematical practice have been thriving. Such approaches include the study of a wide variety of issues concerned with the way mathematics is done, evaluated, and applied, and in addition, or in connection therewith, with historical episodes or traditions, applications, educational problems, cognitive questions, etc. We suggest using the label "philosophy of mathematical practice" as a general term for this gamut of approaches, open to interdisciplinary work. (APMP, 2017)

We observe that in this mission statement of the APMP, philosophy of mathematical practice is not primarily defined in terms of its object, but rather by its methods: it is a term for a "gamut of approaches" applied to the philosophy of mathematics. Löwe (2016, p. 31) reports that

> at the inaugural conference of the *Association for the Philosophy of Mathematical Practice* in Brussels there was a critical discussion of the term "Philosophy of mathematical practice". Its syntactic form "philosophy of *X*" suggests that there is an object "mathematical practice" whose philosophy it is studying. [...] This view was in general rejected by the participants of the inaugural conference; instead, the consensus was that philosophy of mathematical practice is an approach (or a collection of approaches) to philosophy of mathematics.

Other terms than "philosophy of mathematical practice" have been used that make it more apparent that the research community is determined by an approach rather than a subject matter, among

---

[1]For a more detailed discussion of this research community and its relationship to philosophy, cf. (Löwe, 2016, §§ 2&3).

them "empirical philosophy of mathematics" (Löwe et al., 2010), "practice-based philosophy of mathematics" (Dutilh Novaes, 2012) or "(socio-)empirically informed philosophy of mathematics" (Müller-Hill, 2009, 2011).

The "gamut of approaches" invoked by the APMP includes historical approaches, sociological approaches, methods from cognitive science and cognitive psychology, as well as mathematics education. The wide range of this list resulted in a cautionary commentary by Jullien and Soler (2014, p. 228; emphasis in the original) remarking that these are

> not, strictly speaking, approaches 'in the philosophy of mathematical practice' [...] They are, rather, [...] *non*-philosophical perspectives on mathematical practice that are *used* by philosophers of mathematical practice or, more prudently, on which *some* philosophers of mathematical practice *can find* [it] *relevant* to rely.

The approaches used are methods of empirical social research (broadly speaking) which has two consequences. Firstly, philosophers of mathematics are rarely trained in the methods of empirical social research. This lack of expertise requires either extensive training or increased collaboration with researchers from other disciplines (cf. Löwe, 2016, §5). Second, since there is hardly any well established methodology for using empirical results in philosophical arguments, reflection about the methods to be used is a worthwhile first step of this particular philosophical endeavour. We contribute to this reflective process in this paper.[2]

3   Experimental philosophy

"Experimental philosophy is a new interdisciplinary field that uses methods normally associated with psychology to investigate questions normally associated with philosophy. (Knobe et al., 2012, p. 81)" This definition of the field of *Experimental Philosophy* establishes it as a field defined by its method rather than its subject, reminding us of the discussion in §2. Another similarity is that experimental philosophy grew out of a methodological dissatisfaction with the *status quo*: experimental philosophers such as Knobe (2007, p. 72) and Prinz (2008, p. 199) have criticized traditional philosophers for making empirical claims on the basis of rather limited and potentially biased data;[3] they propose to replace expert intuition by the aforementioned "methods normally associated with psychology". While the method of experimental philosophy has been mostly applied in epistemology, philosophy of mind, philosophy of language, and ethics, in principle, there are no obstacles to applying these methods to any subfield of philosophy, including philosophy of mathematics. We observe that the mentioned definition of "experimental philosophy" does not mention the experimental method specifically (cf. §4).[4]

Given these close similarities, it is not surprising that experimental philosophy had to deal with the two issues mentioned at the end of §2: researchers in experimental philosophy had to develop skills in psychological research methods and the field had to reflect on its own methodology.

---

[2]We should like to emphasise that in this paper, we do not intend to contribute to the discussion of whether empirical methods are fruitful in philosophy or the even more overarching question about the aims of philosophy. We discuss the methodological issues under the tacit assumption that we have decided to apply empirical methods in philosophy.

[3]Inglis and Aberdein (2016, p. 166) introduce the term "examplar philosophers" for those who "offer an example [...], assert that [it] has a given property, and appeal to the readers' intuition for agreement".

[4]Since this paper is not chiefly about experimental philosophy, we shall not be able to do justice to the scope and diversity of the experimental philosophy literature that includes much more than what we can discuss in this section.

Concerning the latter point, from its very inception, experimental philosophy was embroiled in a methodological debate with the armchair philosophers it aimed to criticise (cf., e.g., Williamson, 2004; Sosa, 2007; Symons, 2008) and reacted by providing a manifesto (Knobe and Nichols, 2008) with a discussion of the method, its goals, and its limitations. Over the years, experimental philosophers have remained acutely aware and keenly interested in discussions of methods (e.g., Alexander, 2010; Sandis, 2010; Strickland and Suben, 2012; Sytsma and Livengood, 2012; Andow, 2016). As a reaction to the recent replication crisis (cf., e.g., Begley and Ioannidis, 2015; Baker, 2016), the experimental philosophy community reacted by setting up the *XPhi Replicability Project* in order to reach a reliable estimate of the replicability of empirical results in experimental philosophy (the organisers of the project, Florian Cova and Brent Strickland, expect to publish its results in the journal *Review of Philosophy and Psychology*). As a consequence, experimental philosophy is a field with a high degree of methodological reflection and thus, we consider it beneficial for philosophers of mathematical practice to learn from the experiences made by experimental philosophers.

Let us return to the use of the term "experimental" in the name of the field. Until recently, the term *experimental philosophy* would have rather evoked the title of Margaret Cavendish's 1666 *Observations Upon Experimental Philosophy* in the minds of readers. Since in Cavendish's time, science was referred to as "natural philosophy", the term *experimental philosophy* here stands for the empirically based method of the natural sciences that only rather recently had embraced the controlled experiment as its most significant tool:

> The seventeenth century witnessed the rise of experiment as a means of acquiring knowledge about nature. It is not as if experiments had had no role in natural philosophy or medicine before the seventeenth century, but rather that the quantity, nature and significance of early modern experiments were markedly superior to what had preceded them in any period in the history of science. (Anstey, 2014, p. 103)

It is easy to forget that the experimental method which many consider paradigmatic for the scientific method as a whole is a relatively modern addition to the toolbox of the scientist. We use this as an excuse to explore the meaning of the word "experimental" in the subsequent section.

4   What is in a name?

According to Merriam-Webster (2017), an experiment is "an operation or procedure carried out under controlled conditions in order to discover an unknown effect or law, to test or establish a hypothesis, or to illustrate a known law". In general, the model for an experimental set-up has three types of variables: *independent variables* manipulated by the experimenter, *dependent variables* that are being measured in order to determine the effect that the change of the independent variables has, and *extraneous* or *controlled variables*. The treatment of these three types of variables constitutes the experimental method:

> The experimental method is defined by the manipulation of independent variables and the measurement of dependent variables. Extraneous variables are either controlled or allowed to vary randomly. In particular, care is taken to remove any variables that are confounded with the independent variables. (Healey and Proctor, 2003, p. ix)

The term *controlled variables* evokes the image of laboratory control in physical experiments where, e.g., the temperature or the pressure during the experiment is kept constant, but this type of control

is hardly ever possible in the social sciences where control of the extraneous variables is usually achieved statistically by random assignments of test subjects (or, more generally, units of analysis) to different conditions. In this paper, we are using the term *control* to refer to both mentioned types of control.

In textbooks on scientific methodology, *experiment* is usually contrasted with *observation* which is missing the feature of manipulation of the independent variables and the control over the extraneous variables.[5] However, not all definitions of the word *experiment* emphasise all three mentioned aspects in order to contrast experiments and observations; in some cases, only the active manipulation of the independent variables is used as the crucial difference between an experiment and an observation: "To experiment is to isolate, prepare, and manipulate things in hopes of producing epistemically useful evidence" (Bogen, 2017, § 1). In the social science literature, studies that lack control of the extraneous variables via randomised assignments of test subjects to different conditions is often called *quasi-experimental* or *correlational*. It is also interesting to note that the stark contrast between experiment and observation did not always exist; e.g., Anstey (2014, p. 105) reports that "in early modern English, the expression 'observation and experiment' [...] is often used as a hendiadys".

The usage of the words "experiment" and "experimental" are more multi-faceted than any textbook definitions can capture. Let us add one particularly striking example of nomenclature from the experimental philosophy literature: Prinz (2008) proposed to demarcate the boundary line between *experimental* and *empirical* based on whether the empirical work is done by the philosophers themselves or rather by other scientists ('mining the data' vs 'collecting the data'):

> Some philosophers make use of empirical results that have been acquired by professional scientists. [...] These results are used to support or refute philosophical theories. I will call this approach 'empirical philosophy'. Other philosophers also conduct their own psychological experiments, an approach known as 'experimental philosophy'. (Prinz, 2008, p. 196)

Löwe (2016, § 4.3) argues that Prinz's demarcation line according to the person collecting the data is important, but the terms "experimental" and "empirical" for the corresponding types of philosophy are infelicitous.[6]

The confusion about when to properly call a method "experimental" is related to the fact that the idea of a well-defined universal notion of the scientific method is a gross oversimplification.[7] Hoyningen-Huene (2013, p. 4*sq*) reports that

> belief in the existence of scientific methods conceived of as strict rules of procedure has eroded. [...] Research situations [...] are so immensely different from each other across

---

[5]Cf., e.g., (Oehlert, 2010, p. 2): "We are in control of experiments, and having that control allows us to make stronger inferences about the nature of differences that we see in the experiment. This [...] distinguishes an experiment from an observational study. An observational study also has treatments, units, and responses. However, in the observational study we merely observe which units are in which treatment groups; we don't get to control that assignment."

[6]Using a sartorial analogy, Löwe (2016, § 4.3) introduces the terms *ready-to-wear* or *off-the-rack* for the use of empirical results in philosophy where the philosopher had no input in the design of the empirical study, *bespoke* for a project in which the philosopher works very closely with the empirical scientist and designs an experiment or other observational activity jointly with her, and *do-it-yourself* as the extreme case of bespoke where the philosopher becomes an empirical scientist and does the empirical work herself.

[7]Cf. (Andersen and Hepburn, 2016).

the whole range of the sciences and across time that it appears utterly impossible to come up with one set of universally valid methodological rules.

It is not fruitful to get into a nomenclatural debate about the correct definition of the term "experimental". Instead, it is more apt to realise that the diairetic distinctions discussed are important features that allow us to describe particular methodological approaches and discuss their strengths and weaknesses. The distinctions that we have seen so far are:

(i) Does the researcher actively manipulate the environment or not?

(ii) Does the researcher control the features that are currently not under investigation or not?

(iii) Is the philosopher actively involved either in the design of the research or even in the actual empirical research itself or not?

Of course, these three mentioned distinctions are not the only ones that matter. No discussion of methodological distinctions of empirical work is complete without mentioning the divide between *qualitative* and *quantitative* research. The language used in our original discussion of the experimental method (in particular, the use of the word "variable") suggests that an experimental set-up presupposes a quantitative methodology, but this is not the case: questions of whether the researcher should actively manipulate the environment or not feature very prominently in discussions of qualitative research as well.

We summarise: there is a multitude of scientific methods, many ways to classify them in order to separate one method from another, each with its own advantages and disadvantages. We should like to propose the term *empirical philosophy* for the philosophical methodology that takes empirical data, produced according to one of the established empirical scientific methods, into account.[8] Both philosophy of mathematical practice and experimental philosophy discussed in §§ 2&3 fall under this label.

The philosophical methodology of incorporating empirical findings in philosophical arguments is underdeveloped; the methodological reflection of experimental philosophers mentioned in § 3 has started this necessary debate (cf. also Crupi and Hartmann, 2010).

5   "Medicine is not an exact science"

In § 4, we discussed a plethora of empirical methods, of which the experimental method in the strict sense, i.e., involving manipulation of the environment under controlled or randomised conditions, is but one. So, it is striking that the dominant image of scientific activity is that of

hypothesis / experimental design / experiment / validation or refutation

---

[8]As a label, *empirical philosophy* has often been identified with the Vienna Circle conception of what philosophy ought to be like: subordinate to (natural) science (Benjamin, 1939). The label was also used in the edited volume (Wagenknecht et al., 2015) where the focus is mostly on qualitative methods and in their *Prolegomena to an Empirical Philosophy of Science*, Osbeck and Nersessian (2015, pp. 15) explicitly contrast their approach with experimental philosophy: "An example is evident in relation to the recent trend of adopting empirical methods from psychology to inform philosophy, including philosophy of science. This 'experimental philosophy' is something of a curiosity, because nowhere has the ambiguity of 'empirical' created more problems than in the discipline of psychology."

We emphatically do not intend to include these connotations in our notion of empirical philosophy. An alternative label that could be used to avoid the confusion with other uses of the term is *empirically-based philosophy*.

which entirely ignores the methodological variety of scientific practice. The overemphasis of the experimental method in popular accounts of the scientific method stems from the fact that in (parts of) the exact sciences, this particular methodology goes a long way.[9] Lieberson and Lynn (2002) lamented that the perceived equivalence of scientific rigour with the methods of the exact sciences is very harmful for the scientific methodology in the social sciences:

> Our thesis is that deeply ingrained in sociology and other social sciences is a special model of natural science that is exceptionally inappropriate. It is derived from physics, particularly the classical physics that existed before the beginning of the twentieth century. [...] The use of physics as the ideal model for sociology is so embedded in our thinking that the influence and appropriateness of this particular model is rarely questioned. [...] Our thesis is that other natural sciences actually offer epistemological and procedural models that are more relevant for the obstacles encountered in sociology and other social sciences. (Lieberson and Lynn, 2002, p. 2)

This complaint echoes the debates of the *Methodenstreit* of the late 19th century: in the *Methodenstreit*, the historical school of economics, led by Gustav von Schmoller (1838–1917), claimed that economic laws could not be independent of historical developments, and that economics should above all proceed on historical-empirical (inductive) grounds; in contrast, Carl Menger (1840–1921) argued that economics should adhere to the hypothetico-deductive methodology of the exact sciences.[10] In some respects, the *Methodenstreit* has never seen a truce.

A similar contrast with the method of the exact sciences is present in the oft-uttered adage "medicine is not an exact science". The phrase originated in the late 19th century in discussions of medical and forensic experts warning against overconfidence in the precision and definiteness of the statements of medical forensics.[11] The intended connotation of the phrase "medicine is not an exact science" is closely related to our earlier discussion of the experimental method: as discussed in §4, the main ingredients of the experimental method in the strict sense are manipulation of the independent variables and control of the extraneous variables: this usually requires an experimental setting in a laboratory, in which the experimenter can "isolate, prepare, and manipulate". Even in the exact sciences, it is questionable whether a researcher can truly control all relevant parameters.[12] But the field of (clinical) medicine usually deals with an individual living human being some of whose variables certainly cannot and some should not be experimentally controlled and no random variation is possible. As a consequence, in the terminology used in §4, medical practice is observational rather than experimental. When medical research is experimental, i.e., done under controlled conditions in a laboratory *in vitro*, the applicability of these research results often suffers from the fact that they were obtained in controlled laboratory environments, but applied to human beings *in vivo*, embedded in a social environment. This phenomenon is witnessed by contradictory

---

[9]We take the liberty of not defining the term "exact sciences", not because we like to stay vague and suggestive, but because this would open yet another subtle terminological debate of which this paper already has its fair share. If the reader is vexed by this, she or he can replace the term "exact sciences" with "classical physics" throughout. We should, however, like to emphasise that even in the prototypical exact science, physics, not all theories lend themselves to experimental testing in the proper sense: e.g., you cannot do controlled experiments concerning planetary motions.

[10]For details, cf. (Huff, 1984, pp. 28–34) and (Schulak and Unterköfler, 2011, pp. 21–28).

[11]Cf. (Mnookin, 2007). Recently, a new wave of overconfidence of this type was dubbed the *CSI effect* and studied in (Shelton et al., 2006; Cole and Dioso-Villa, 2009; Smith et al., 2011; Cole, 2015).

[12]We will not go into details here and refer the reader to the rich literature on *ceteris paribus* clauses in philosophy of science (cf.,e.g., Earman et al., 2003; Schrenk, 2007).

newspaper headlines of the type "Red wine is good for your health" and "Red wine is not good for you after all".

But it is not just that controlling variables is harder and possibly unethical in the medical sciences; also the concepts that medicine is dealing with are noticeably different from the concepts in physics. E.g., the concept of "good for you" is normative, and therefore, its relevant standards are clearly determined by social, societal, and psychological norms in a way that concepts in the exact sciences usually are not. Notice that we do not take a position on whether the concepts in the natural sciences are socially constructed.[13] But while in the natural sciences, it is a respectable philosophical position to claim that the concepts of *force* or *gravity* are entirely independent of human culture and society, it would be bizarre to claim that a normative concept such as "good for you" could be defined in full generality independently of references to human psychology and society. When moving from medical sciences to the social sciences, this becomes even more patently obvious.

The crucial step in empirical social research to deal with the social dependence of the concepts is that of *operationalisation*: the researcher models the socially dependent concept with an operationalised concept for which—within the confines of the given study—it can be determined with a reasonable degree of certainty whether the concept applies or not. In other words, the concepts we are dealing with are either *subjective*, *vague*, or *cluster concepts*[14] and operationalisation replaces these concepts with a precisely defined concept.

Let us consider an example: if a political scientist aims to study the behaviour of *typical Conservative Party voters*, she will need to operationalise this concept in order to have a precise assignment of the label *typical Conservative Party voter* to individual test subjects. E.g., based on the past voting records, she counts a test subject as *typical Conservative Party voter* if they voted for the Conservative Party at least three times out of five. Note that this is not a definition of a *typical Conservative Party voter*: the researcher is very much aware that, e.g., there might be typical Conservative Party voters who, for whatever reason, did not vote for the Conservative Party in the last few elections, creating a false negative (or similarly for a false positive). But as there is no precise definition, its role has to be played by the operationalised concept. This, in turn, requires arguments that this particular operationalisation (producing false positives and negatives) does not affect the results of this study. The choice of operationalisation and the arguments for it are confined to the concrete study that the political scientist is working on.[15]

In the context of the philosophical debate, the need to operationalise is closely connected to the distinction between data and phenomena highlighted by Bogen and Woodward (1988, pp. 305*sq*):

> Data, which play the role of evidence for the existence of phenomena, for the most part can be straightforwardly observed. However, data typically cannot be predicted or systematically explained by theory. By contrast, well-developed scientific theories do predict and explain facts about phenomena. Phenomena are detected through the use of data, but in most cases are not observable in any interesting sense of that term.

In our example, while the data, i.e., the voting record and the individual behaviour, can be straightforwardly observed, the phenomenon that the political scientist aims to study cannot be read off

---

[13]We refer the reader to the relevant literature (e.g., Longino, 2016).

[14]I.e., concepts defined by a list of criteria none of which are necessary or sufficient, akin to the Wittgensteinian *Familienähnlichkeit* (Gasking, 1960; Cooper, 1972; Parsons, 1973).

[15]Note that this distinguishes operationalisation from Carnapian *explication*.

directly from the data, but requires conceptual consolidation into a phenomenology.[16] The replacement of the concept from the original research question with the operationalised concept is one of the underlying reasons for criticism of the use of empirical techniques in philosophy: e.g., Kauppinen (2007, p. 97) argues that experimental studies about the concept of knowledge only study the *folk concept* of knowledge which may not be related to the philosophical concept that—as philosophers—we aim to study. As said before, the onus is on the empirical researcher to argue that the replacement of the informal concept from the original question with the operationalised concept is adequate within the context of the study.

All of the above issues combined result in a situation where the empirical social scientists need to remain aware that any question they are asking is not universal, but restricted to a particular cultural and social context, and that data collected from arbitrary sources is most likely meaningless. Kauppinen (2007, pp. 101*sq*) reminds us that

> [i]t should be obvious that when philosophers appeal to 'us' in making their claims, the extension is limited to those who are competent with the concept in question. After all, what incompetent users of a concept say about a given case does not tell us anything about the concept we are interested in—someone who has no relevant pre-theoretical knowledge about the concept cannot manifest it. Nobody would test a Gettier analysis by asking a small child whether the person in the case described knows or not, or count the child's response as a counterexample. And children are only the most obvious example. On many theories of concept possession, competence with a concept is a matter of degree and context. This is to deny that there is, strictly speaking, such a thing as a 'competent speaker of English', for example.

It is important to note that in practice, determining whether a given test subject has the competence to answer our question is not isolated from and independent of the answer they give; this raises the spectre of circular reasoning and leading to meta-phenomena such as Collins's *experimenter's regress* (Collins, 1981). Empirical social research is fundamentally hermeneutic in nature: traditionally, hermeneutics is the skill of interpreting *oratio obscura* by taking commonsense reasoning, the context, and the background of the speaker into account; it is governed by principles such as presumptions and *aequitas hermeneutica* (Scholz, 1999, § I.7–10). Thus, the hermeneutic approach is in conflict with some of the fundamental principles of the experimental approach which requires the control of external factors, whereas the hermeneutic principles require the *consideration* of the effects of many of these external factors as an important part of the interpretation. In Dilthey's terms, the social sciences are about *Verstehen* rather than *Erklären*.[17]

Taken together, the impossibility of controlling variables and the ubiquity of cluster concepts that need to be replaced by an operationalisation in the context of empirical studies, requires an approach to the empirical social sciences that includes multiple viewpoints, angles and methods. Tilly (2004, p. 597) summarises the situation as follows:

---

[16]Löwe and Müller (2011) discuss this in the framework of *conceptual modelling* and use the distinction to shed some light on the methodological discussion between the experimental philosophers and the armchair philosophers. They note that the instability of introspective intuitions has been criticized long before the advent of experimental philosophy: "[T]he intuitive findings of different people, even of different experts, are often inconsistent. [...] If agreement about usage cannot be reached within so restricted a sample as the class of Oxford Professors of Philosophy, what are the prospects when the sample is enlarged? (Mates, 1958, p. 165)"

[17]Cf. also the discussion of the *That's-Not-All-There-Is Objection* against experimental philosophy (Knobe and Nichols, 2008, p. 10).

> [a] common prejudice, to be sure, divides the social world into phenomena that are suitable for quantification (population distributions, social mobility, etc.) and those that are irreducibly qualitative: conversation, narratives, biography, ethnography, and history often serve as examples. Formalisms clearly can and do apply, however, to these phenomena as well. [. . .] Formalisms blindly followed induce blindness. Intelligently adopted, however, they improve vision. Being obliged to spell out the argument, to check its logical implications, and to examine whether the evidence conforms to the argument promotes both visual acuity and intellectual responsibility.

The meta-method used for obtaining this "visual acuity and intellectual responsibility" is that of *methodological triangulation* which we shall discuss in the next section.

6 Triangulation

The method of *triangulation*, sometimes called *mixed method research*, is a (meta-)methodology in social research. The use of the term can be traced to a technical paper focusing on different quantitative methods by Campbell and Fiske (1959).[18] The seminal book (Webb et al., 1966, p. 3) generalised the idea to the general notion of combination of multiple methods of data collection:

> The most persuasive evidence comes through a triangulation of measurement processes. If a proposition can survive the onslaught of a series of imperfect measures, with all their irrelevant error, confidence should be placed in it. Of course, this confidence is increased by minimizing error in each instrument and by a reasonable belief in the different and divergent effects of the sources of error.

This meta-method of combining methodologies in the study of the same phenomenon was taken up by Denzin (1970, 1978) in his influential textbook.[19]

Triangulation allows the empirical researcher to avoid the danger of circularity implicit in the hermeneutic approach: e.g., if determining the competence of a test subject and the subject's answers to the questionnaire are closely related and depend on each other, then a solution is to use

---

[18]In (Campbell and Fiske, 1959), the term refers to actual triangles in tables of data. Nowadays, we usually think of the triangulation metaphor as mimicking geometrical triangulation in navigation, allowing to determine the location of an object by giving its coordinates from two different angles. This notion of triangulation is not to be confused with the notion used by Davidson (1982) which uses the same metaphor in a similar situation.

[19]Denzin (1978) also gives a classification of types of triangulation: data triangulation, investigator triangulation (involving multiple researchers to avoid biases brought in by the researcher), theory triangulation (using multiple theoretical schemes), methodological triangulation. We focus on methodological triangulation for the present discussion without claiming that the other types of triangulation are unimportant.

Green and Wortham (2018, p. 69–72) argue that W. E. B. Du Bois (1868–1963), the first African American to receive a doctorate from Harvard University, was using methodological triangulation much earlier than this, and give examples from studies performed by Du Bois and published between 1898 and 1904: "Throughout his research, Du Bois used a diverse array of research methods to gather data. ... Few, if any sociologists of the era, were as imbued with a methodological conscience as the intrepid Du Bois (Green and Wortham, 2018, p. 71)". (Cf. also Heesen et al., 20xx, § 2.)

Du Bois's cautious and critical attitude towards the research methods is best described by his notion of *diffidence* laid out in (Du Bois, 1899, pp. 2–3): "The best available methods of sociological research are at present so liable to inaccuracies that the careful student discloses the results of individual research with diffidence; he knows that they are liable to error from the seemingly ineradicable faults of the statistical method, to even greater error from the methods of general observation, and, above all, he must ever tremble lest some personal bias, some moral conviction or some unconscious trend of thought due to previous training, has to a degree distorted the picture in his view."

one method to determine competence of the test subject and an entirely different method to do the survey; or, multiple different operationalisations of an informal concept could be used in order to provide an argument that the answer does not depend on the operationalisation.[20] Empirical results obtain stability if several methodological approaches converge on them and only in this way can social phenomena be sufficiently stabilized in order to draw reliable conclusions. Following Geertz (1973), the results of such multi-modal approaches are often called *thick descriptions*.[21]

Even though triangulation has been an accepted meta-method in the social sciences for decades, a large amount of social empirical research is done using a single method, often an experimental method. It has been argued that this state of affairs is one of the reasons why the social sciences, and in particular psychology, have been hit especially hard by the replication crisis (Stroebe and Strack, 2014; Earp and Trafimow, 2015); one of the proposed solutions to the replication crisis is a form of triangulation (cf. Schooler, 2014, there called *meta-science*).[22]

We propose to embrace and apply the meta-methodology of triangulation for what we called *empirical philosophy* in § 4. We should hasten to add that this proposal does not mean that we think that any empirical approach that uses but one method is flawed and without merit. In fact, as can be expected from the standard popular account of falsificationalist science, the experimental approach is particularly good at producing negative results and casting doubt on hypothesis and premises. This critical approach to hidden premises is very much needed in philosophy; e.g., the results of Weinberg et al. (2001), without any further triangulation by other methods, are sufficient to cast doubt on the assumption that epistemic intuitions are universal traits of human beings independent of which culture they grew up in: consequently, if a philosophical position is based on this assumption, the results require any philosophers who holds these views to re-evaluate their position.

7   Triangulation in philosophy of mathematical practice

After our excursion into general methodology or philosophy of science, we now return to the philosophy of mathematical practice: we apply the lessons learned from considering methodological issues in the empirical social sciences. As in the general case in § 6, we emphasise that not all empirically-based philosophical analysis that relies on a single method is problematic: the experimental method is very good at pointing out hidden assumptions about homogeneity or universality and casting doubt on these. E.g., Núñez (2011) uses an experimental (or quasi-experimental) approach and without any further triangulation by other methods, his results show convincingly that there are problems with the assumption that previous experiments prove that the cognitive number line is innate and physically represented in the human brain.

However, raising issues with assumptions is not all that philosophy of mathematics wants to do,

---

[20]In Denzin's terminology of Footnote 19, the former example would be methodological triangulation, the latter would be theory triangulation.

[21]It is interesting to note that Collins's own solution to the experimenter's regress problem is a form of unintentional triangulation; in his view, "the regress is eventually broken by negotiation within the appropriate scientific community, a process driven by factors such as the career, social, and cognitive interests of the scientists, and the perceived utility for future work (Franklin and Perovic, 2016, § 1.2.1)".

[22]Heesen et al. (20xx) propose a formal model to evaluate and discuss the merits of triangulation based on Du Bois's *diffidence* (cf. Footnote 19): they discuss criticism of methodological triangulation in the literature (e.g., that it produces "discordant evidence") and conclude that their model "vindicate[s the] ... use of methodological triangulation (Heesen et al., 20xx, § 4)".

and in analogy to the use of triangulation in the empirical social sciences, we propose to embrace methodological triangulation as a meta-method in empirical philosophy of mathematics and close our paper by giving three examples of good triangulation practice in philosophy of mathematics:

Our first example is the study on knowledge attributions among mathematicians. Müller-Hill (2009) discusses a questionnaire study in the tradition of experimental philosophy: test subjects from the target group (research mathematicians, defined as people who claim to have research or university teaching experience in mathematics) were given questionnaires with fictitious stories about the production of mathematical knowledge and the resulting data were statistically evaluated (cf. also Löwe et al., 2010). In one of the stories, a fictitious mathematician believes in the truth of a theorem (the proof of the *Jones conjecture*) he published, but finds out during a talk he attends that there is a counterexample to his result. 61.3% of the test subjects answer positively to the question "Does John know that the Jones conjecture is false?" and yet 71% of the test subjects answer positively to the question "Did John know that the Jones conjecture was true on the morning before the talk?". As in the methodological debates about experimental philosophy, traditional philosophers of mathematics have criticised questionnaire-based studies like this with arguments similar to the objections listed in Knobe and Nichols (2008): in the terminology used in § 5, traditional philosophers observe that such a study replaces the notion of mathematical knowledge with an operationalised concept that is based on actual knowledge utterances by research mathematicians and claim that this operationalised concept is quite different from the philosophical concept of mathematical knowledge.

Müller-Hill (2011) then interpreted the quantitative results of the initial study by means of frequency and cluster analysis and supplements this with the text comments provided by the test subjects (Müller-Hill, 2011, § 3.2.2). The qualitative processing of the various answers in the initial study allowed Müller-Hill to isolate a number of different answer profiles for each fictitious scenario. These then formed the basis of a number of in-depth interviews, serving as a qualitative complement to the initial study and thus an additional triangulation point (Müller-Hill, 2011, § 4).

The second example are studies by the research groups at Loughborough University and Rutgers University of the conviction of mathematicians about the correctness of a mathematical result. The degree of conviction about mathematical correctness is one of those highly context-dependent concepts that defy precise definitions and thus need to be operationalised. In a preliminary study, Inglis and Mejia-Ramos (2008) provided a typology of responses to the question about the level of persuasion by a proof. Weber (2008) and Weber and Mejia-Ramos (2011) conducted interview studies on goals and methods of reading proofs by research mathematicians; these qualitative studies generated hypotheses that were then tested with a survey study with 118 research-active mathematicians by Mejia-Ramos and Weber (2014). Further focusing on the mechanisms of reading proofs, an eye-tracking study of mathematicians reading proofs by Inglis and Alcock (2012) produced data that was provided to Weber and Mejia-Ramos (2013) who then further analysed the data with other methods; their study was followed up by a reply of Inglis and Alcock (2013).

This example is of particular interest since it illustrates how triangulation can result in what Stegenga (2012) calls "discordant evidence": on some particular issues relevant to the study, the two research groups reached different conclusions using different methods. On the basis of introspective data gathered through interviews and questionnaires, Weber and Mejia-Ramos claimed that mathematicians typically skim mathematical texts before reading them in detail, but the eye-tracking study by Inglis and Alcock suggested the opposite. Such instances of scientific dissent fuelled by methodological pluralism are an opportunity and invitation to renewed and intensified

discussion, also on the methodological level, that one would hope leads eventually to a more nuanced understanding of the phenomena.

Our final example is the study of the mathematical peer review process: the original epistemological motivation of Geist et al. (2010) was the question about mathematical certainty in the case of knowledge by testimony via published proofs. Geist et al. (2010) studied the role of the mathematical peer review process in ascertaining correctness of proofs in the published literature, in particular, whether the fact that a paper went through the mathematical peer review process serves as a warrant for mathematical correctness. The idealised description of the peer review process would assume that every referee meticulously checks the mathematical details of proofs, but that idealised description does not match with the experiences of many if not most mathematicians. In the absence of empirical studies of the mathematical peer review process, Geist et al. (2010) systematically assess the process via a (small-scale) qualitative questionnaire study of editors of mathematical journals and a study of referee agreement comparing mathematical conferences with conferences from other subjects. Other researchers arrived at similar empirical questions from very different theoretical backgrounds: both Weintraub and Gayer (2001) and Heinze (2010) are sociologically interested in the criteria used by the community of mathematicians to decide whether a new proof becomes accepted; Weintraub and Gayer study this via an in-depth analysis of one particular case (the proof of the existence of a competitive equilibrium by Arrow and Debreu) and Heinze investigates this by means of a questionnaire study. Recently, the sociologist Greiffenhagen (personal communication) started using interviews with journal editors and case studies dealing with similar questions.

The last example is different from the other two in that it is not the same research group approaching the question from different angles, but different unrelated groups, in this case from three different disciplines (philosophy, mathematics education, and sociology) coalesce to the same questions from very different starting points with possibly incompatible motivations. Triangulation in the philosophy of mathematics, i.e., approaching philosophical questions empirically from multiple angles and with multiple methods should not ultimately rely on serendipitous convergence of interests: in order to increase the number of triangulated studies in empirical philosophy of mathematics, more dialogue among reseachers is needed to overcome this motivational incommensurabilities and find common ground (as argued by Löwe, 2016, p. 39). We hope that this paper will convince some of our colleages to commit some of their energy to this worthwhile effort.

References

Alexander, J. (2010). Is experimental philosophy philosophically significant? *Philosophical Psychology*, 23(3):377–389.

Andersen, H. and Hepburn, B. (2016). Scientific method. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2016 edition.

Andow, J. (2016). Qualitative tools and experimental philosophy. *Philosophical Psychology*, 29(8):1128–1141.

Anstey, P. R. (2014). Philosophy of experiment in early modern England: The case of Bacon, Boyle and Hooke. *Early Science and Medicine*, 19:103–132.

APMP (2017). Association for the Philosophy of Mathematical Practice. About. Website accessed on 4 October 2017.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604):452–454.

Begley, C. G. and Ioannidis, J. P. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research*, 116(1):116–126.

Benjamin, A. C. (1939). What is empirical philosophy? *Journal of Philosophy*, 36(19):517–525.

Bogen, J. (2017). Theory and observation in science. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2017 Edition.

Bogen, J. and Woodward, J. (1988). Saving the phenomena. *Philosophical Review*, 97(3):303–352.

Campbell, D. and Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2):81–105.

Cole, S. A. (2015). A surfeit of science: The "CSI effect" and the media appropriation of the public understanding of science. *Public Understanding of Science*, 24(2):130–146.

Cole, S. A. and Dioso-Villa, R. (2009). Investigating the "CSI effect" effect: Media and litigation crisis in criminal law. *Stanford Law Review*, 61(6):1335–1373.

Collins, H. M. (1981). 'Son of seven sexes', the social destruction of a physical phenomenon. *Social Studies of Science*, 11(1):33–62.

Cooper, D. E. (1972). Definitions and 'clusters'. *Mind*, 81(324):495–503.

Corfield, D. (2003). *Towards a Philosophy of Real Mathematics*. Cambridge University Press.

Crupi, V. and Hartmann, S. (2010). Formal and empirical methods in philosophy of science. In Stadler, F., Dieks, D., González, W. J., Hartmann, S., Uebel, T., and Weber, M., editors, *The Present Situation in the Philosophy of Science*, volume 1 of *The Philosophy of Science in a European Perspective*, pages 87–98. Springer-Verlag.

Davidson, D. (1982). Rational animals. *Dialectica*, 36:318–327.

Denzin, N. K. (1970). *The research act: A theoretical introduction to sociological methods*. Aldine, 1st edition.

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. McGraw-Hill, 2nd edition.

Du Bois, W. E. B. (1899). *The Philadelphia Negro: A social study*. University of Pennsylvania Press.

Dutilh Novaes, C. (2012). Towards a practice-based philosophy of logic: formal languages as a case study. *Philosophia Scientiae*, 16:71–102.

Earman, J., Glymour, C., and Mitchell, S., editors (2003). *Ceteris Paribus Laws*. Springer-Verlag.

Earp, B. D. and Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, 6. Article 621.

Franklin, A. and Perovic, S. (2016). Experiment in physics. In Zalta, E., editor, *The Stanford Encyclopedia of Philosophy*. Winter 2016 Edition.

Gasking, D. (1960). Clusters. *Australasian Journal of Philosophy*, 38(1):1–36.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In Geertz, C., editor, *The Interpretation of Cultures. Selected Essays*, pages 3–30. Basic Books.

Geist, C., Löwe, B., and Van Kerkhove, B. (2010). Peer review and knowledge by testimony in mathematics. In Löwe, B. and Müller, T., editors, *PhiMSAMP, Philosophy of Mathematics: Sociological Aspects and Mathematical Practice*, volume 11 of *Texts in Philosophy*, pages 155–178. College Publications.

Green, D. S. and Wortham, R. A. (2018). The sociological insight of W. E. B. Du Bois. *Sociological Inquiry*, 88(1):56–78.

Healey, A. F. and Proctor, R. W., editors (2003). *Handbook of Psychology. Volume 4: Experimental Psychology*. John Wiley & Sons.

Heesen, R., Kofi Bright, L., and Zucker, A. (20xx). Vindicating methodological triangulation. to appear in *Synthese*.

Heinze, A. (2010). Mathematicians individual criteria for accepting theorems and proofs: An empirical approach. In Hanna, G., Jahnke, H. N., and Pulte, H., editors, *Explanation and Proof in Mathematics. Philosophical and Educational Perspectives*, pages 101–111. Springer-Verlag.

Hersh, R. (1997). *What is mathematics, really?* Oxford University Press.

Hoyningen-Huene, P. (2013). *Systematicity. The Nature of Science*. Oxford Studies in Philosophy of Science. Oxford University Press.

Huff, T. E. (1984). *Max Weber and the Methodology of the Social Sciences*. Transaction Books.

Inglis, M. and Aberdein, A. (2016). Diversity in proof appraisal. In Larvor, B., editor, *Mathematical Cultures. The London Meetings 2012–2014*, Trends in the History of Science, pages 163–179. Birkhäuser.

Inglis, M. and Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43(4):358–390.

Inglis, M. and Alcock, L. (2013). Skimming: a response to Weber and Mejía-Ramos. *Journal for Research in Mathematics Education*, 44(2):471–474.

Inglis, M. and Mejia-Ramos, J. P. (2008). How persuaded are you? A typology of responses. *Research in Mathematics Education*, 10:119–133.

Jullien, C. and Soler, L. (2014). Conceptions of mathematical practices: Some remarks. Commentary on "The impact of the philosophy of mathematical practice on the philosophy of mathematics", by Jean Paul Van Bendegem. In Soler, L., Zwart, S., Lynch, M., and Israel-Jost, V., editors, *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*, Routledge Studies in the Philosophy of Science, pages 227–237. Routledge.

Kauppinen, A. (2007). The rise and fall of experimental philosophy. *Philosophical Explorations*, 10(2):95–118.

Knobe, J. (2007). Experimental philosophy. *The Philosophers' Magazine*, 50:72–73.

Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkissian, H., and Sommers, T. (2012). Experimental philosophy. *Annual Review of Psychology*, 63:81–99.

Knobe, J. and Nichols, S. (2008). An experimental philosophy manifesto. In Knobe, J. and Nichols, S., editors, *Experimental Philosophy*, pages 3–16. Oxford University Press.

Lieberson, S. and Lynn, F. B. (2002). Barking up the wrong branch: Scientific alternatives to the current model of sociological science. *Annual Review of Sociology*, 28:1–19.

Longino, H. (2016). The social dimensions of scientific knowledge. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Spring 2016 Edition.

Löwe, B. (2016). Philosophy or not? The study of cultures and practices of mathematics. Selected papers from the conference in Guangzhou, China, 9–12 November 2012. In Ju, S., Löwe, B., Müller, T., and Xie, Y., editors, *Cultures of Mathematics and Logic*, Trends in the History of Science, pages 23–42. Birkhäuser.

Löwe, B. and Müller, T. (2011). Data and phenomena in conceptual modelling. *Synthese*, 182(1):131–148.

Löwe, B., Müller, T., and Müller-Hill, E. (2010). Mathematical knowledge: a case study in empirical philosophy of mathematics. In Van Kerkhove, B., De Vuyst, J., and Van Bendegem, J. P., editors, *Philosophical Perspectives on Mathematical Practice*, volume 12 of *Texts in Philosophy*, pages 185–203. College Publications.

Mancosu, P., editor (2008). *The Philosophy of Mathematical Practice*. Oxford University Press.

Mates, B. (1958). On the verification of statements about ordinary language. *Inquiry*, 1:161–171.

Mejia-Ramos, J. P. and Weber, K. (2014). How and why mathematicians read proofs: Further evidence from a survey study. *Educational Studies in Mathematics*, 85:161–173.

Merriam-Webster (2017). Experiment. In *Online Dictionary*. Website accessed on 4 October 2017.

Mnookin, J. L. (2007). Idealizing science and demonizing experts: An intellectual history of expert evidence. *Villanova Law Review*, 52(4):763–801.

Müller-Hill, E. (2009). Formalizability and knowledge ascriptions in mathematical practice. *Philosophia Scientiae*, 13(2):21–43.

Müller-Hill, E. (2011). *Die epistemische Rolle formalisierbarer mathematischer Beweise— Formalisierbarkeitsorientierte Konzeptionen mathematischen Wissens und mathematischer Rechtfertigung innerhalb einer sozio-empirisch informierten Erkenntnistheorie der Mathematik.* PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn.

Núñez, R. (2011). No innate number line in the human brain. *Journal of Cross-Cultural Psychology*, 45(4):651–68.

Oehlert, G. W. (2010). *A First Course in Design and Analysis of Experiments.* Author (Gary W. Oehlert).

Osbeck, L. M. and Nersessian, N. J. (2015). Prolegomena to an empirical philosophy of science. In Wagenknecht, S., Nersessian, N. J., and Andersen, H., editors, *Empirical philosophy of science: Introducing qualitative methods into philosophy of science*, volume 21 of *Studies in Applied Philosophy, Epistemology and Rational Ethics*, pages 13–35. Springer-Verlag.

Parsons, K. P. (1973). Three concepts of clusters. *Philosophy and Phenomenological Research*, 33(4):514–523.

Prinz, J. J. (2008). Empirical philosophy and experimental philosophy. In Knobe, J. and Nichols, S., editors, *Experimental Philosophy*, pages 198–208. Oxford University Press.

Sandis, C. (2010). The experimental turn and ordinary language. *Essays in Philosophy*, 11(2):181–196.

Scholz, O. R. (1999). *Verstehen und Rationalität. Untersuchungen zu den Grundlagen von Hermeneutik und Sprachphilosophie*, volume 76 of *Philosophische Abhandlungen*. Vittorio Klostermann.

Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. *Nature*, 515(7525).

Schrenk, M. (2007). *The Metaphysics of Ceteris Paribus Laws*, volume 16 of *Philosophische Analyse*. Ontos.

Schulak, E. M. and Unterköfler, H. (2011). *The Austrian School of Economics. A History of Its Ideas, Ambassadors, and Institutions.* Ludwig von Mises Institute.

Shelton, D. E., Kim, Y. S., and Barak, G. (2006). A study of juror expectations and demands concerning scientific evidence: Does the "CSI effect" exist? *Vanderbilt Journal of Entertainment and Technology Law*, 9:331–368.

Smith, S. M., Stinson, V., and Patry, M. W. (2011). Fact or fiction? the myth and reality of the CSI effect. *Court Review*, 47:4–7.

Sosa, E. (2007). Experimental philosophy and philosophical intuition. *Philosophical Studies*, 132:99–107.

Stegenga, J. (2012). Rerum concordia discors: Robustness and discordant multimodal evidence. In Soler, L., Trizio, E., Nickles, T., and Wimsatt, W., editors, *Characterizing the robustness of science: After the practice turn in philosophy of science*, volume 292 of *Boston Studies in the Philosophy of Science*, pages 207–226. Springer-Verlag.

Strickland, B. and Suben, A. (2012). Experimenter philosophy: the problem of experimenter bias in experimental philosophy. *Review of Philosophy and Psychology*, 3:457–467.

Stroebe, W. and Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, 9(1):59–71.

Symons, J. (2008). Intuition and philosophical methodology. *Axiomathes*, 18:67–89.

Sytsma, J. and Livengood, J. (2012). Experimental philosophy and philosophical disputes. *Essays in Philosophy*, 13(1):145–160.

Tilly, C. (2004). Observations of social processes and their formal representations. *Sociological Theory*, 22(4):595–602.

Van Kerkhove, B., editor (2008). *New Perspectives On Mathematical Practices. Essays in Philosophy and History of Mathematics. Brussels, Belgium, 26–28 March 2007*. World Scientific.

Van Kerkhove, B., De Vuyst, J., and Van Bendegem, J. P., editors (2010). *Philosophical Perspectives on Mathematical Practice*, volume 12 of *Texts in Philosophy*. College Publications.

Van Kerkhove, B. and Van Bendegem, J. P., editors (2007). *Perspectives on Mathematical Practices. Bringing Together Philosophy of Mathematics, Sociology of Mathematics, and Mathematics Education*, volume 5 of *Logic, Epistemology, and the Unity of Science*. Springer-Verlag.

Wagenknecht, S., Nersessian, N. J., and Andersen, H., editors (2015). *Empirical philosophy of science: Introducing qualitative methods into philosophy of science*, volume 21 of *Studies in Applied Philosophy, Epistemology and Rational Ethics*. Springer-Verlag.

Webb, E. J., Campbell, D. T., Schwartz, R. D., and Sechrest, L. (1966). *Unobtrusive Measures. Nonreactive Research in the Social Sciences*. Rand McNally & Company.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, 39:431–459.

Weber, K. and Mejia-Ramos, J. (2013). On mathematicians' proof skimming. a reply to Inglis and Alcock. *Journal for Research in Mathematics Education*, 44(2):464–471.

Weber, K. and Mejia-Ramos, J. P. (2011). How and why mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics*, 76:329–344.

Weinberg, J. M., Nichols, S., and Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29:429–460.

Weintraub, E. R. and Gayer, T. (2001). Equilibrium proofmaking. *Journal of the History of Economic Thought*, 23(4):421–442.

Williamson, T. (2004). Philosophical 'intuitions' and scepticism about judgment. *Dialectica*, 58(1):109–153.