EXPERIMENTAL PHILOSOPHY, RATIONALISM, AND NATURALISM

RETHINKING PHILOSOPHICAL METHOD

EDITED BY
EUGEN FISCHER AND
JOHN COLLINS

Contents

List of contributors

Unless indicated otherwise, contributors belong to philosophy departments

John Collins, *University of East Anglia, UK*

Tim Crane, *University of Cambridge, UK*

Paul E. Engelhardt, *School of Psychology, University of East Anglia, UK*

Eugen Fischer, *University of East Anglia, UK*

Mikkel Gerken, *University of Edinburgh, UK*

Aurélie Herbelot, *Computer Laboratory, University of Cambridge, UK*

Amir Horowitz, *Open University of Israel*

Hilary Kornblith, *University of Massachusetts, Amherst, USA*

Edouard Machery, *University of Pittsburgh, USA*

Jennifer Nado, *Lingnan University, Hong Kong*

Bence Nanay, *University of Antwerp, Belgium, and University of Cambridge, UK*

David Papineau, *King's College London, UK*

Daniele Sgaravatti, *Roma Tre University, Italy*

Jonathan M. Weinberg, *University of Arizona, USA*

Acknowledgements

Introduction

Rationalism and naturalism in the age of experimental philosophy

Eugen Fischer and John Collins

Much of the excitement in contemporary analytic philosophy arises from many of its major debates being animated by deep disagreements about the nature of philosophy itself, including its actual and proper methods, and its plausible aims and ambitions. After decades of relative neglect, these metaphilosophical issues have, over the last ten-to-fifteen years, not merely become a focus of debate in their own right, but have also shaped current and ongoing discussion of first-order philosophical questions in a range of areas including ethics, epistemology and metaphysics, the philosophy of language, and the philosophy of mind and action.

While in many ways fresh, these debates are driven by a venerable ambition that already had a pedigree when articulated by Kant: the ambition to place philosophy on the secure path of a science (*Wissenschaft*), with 'procedure' or methods that go beyond 'merely random groping', let alone 'groping among mere concepts' (Kant 1787/2003, 21 (B xv)), and which allow philosophers to overcome that curiously persistent disagreement among 'the most excellent minds' that had already exasperated Descartes (1637/1993, 5 (AT 8)), over a hundred years earlier. The ambition is as old as the divisions between philosophers who look for the secure path: Methodological rationalists of different stripes pursue it by trying to develop and defend a priori methods, which draw upon intuition or pure reflection alone, and so hold out the promise of an autonomous philosophy that seeks no warrant or guidance from empirical inquiry. Methodological naturalists explore different ways of addressing philosophical problems by drawing

on a posteriori methods and findings from science.[1] This age-old and continually evolving divide has recently been radically reshaped—once again—through the advent of experimental philosophy.

This philosophical movement—too varied in its aims and methods to qualify as a 'school' or 'approach', let alone 'position'—attempts to employ empirical methods and findings from the social sciences to address philosophical questions and problems. It builds on the assumption that, for better or worse, intuitions are crucially involved in philosophical work. For example, many (but certainly not all) philosophical paradoxes and problems are engendered by intuitions at odds with background beliefs (or among each other), and many analytic philosophers treat intuitions as evidence for or against claims and theories that answer philosophical questions. Experimental philosophers use empirical surveys and experiments to develop an understanding of philosophically relevant intuitions that helps us determine whether we should accept or reject them.[2] While the first generation of experimental philosophy studies focused on the use of survey methods to elicit philosophically relevant intuitions and study their sensitivity to different parameters, a second generation of such studies has come to deploy experimental methods and findings from cognitive and social psychology to develop explanations of such intuitions that facilitate the assessment of our warrant for accepting them (see below, Section 4, pp. 20–23). These efforts have extended significantly beyond card-carrying members of the experimental philosophy movement. They have been, and are, used both to attack and to defend rationalist reliance on intuition and a priori methods—or simply to chart, more precisely, which intuitions philosophers may rely on under which circumstances, and when and where they should beware. Hence, far from overcoming or deciding the debate between methodological rationalism and naturalism, experimental philosophy has recapitulated the division.

The efforts of experimental philosophy, however, have transformed both sides. First, they have exerted pressure on defenders of 'armchair' methods to take into account scientific

findings about the cognitive processes that generate intuitions and facilitate the acquisition of knowledge, a priori and other. Second, they have given methodological naturalism a new twist: in a naturalist vein, they put empirical scientific methods to the service of addressing philosophical problems. But where traditional naturalists sought to address philosophical problems about a topic X (intentional action, consciousness, knowledge, perception, etc.) by building on scientific work about X, most strands of experimental philosophy proceed by building on scientific—namely, psychological—work about the ways in which we think about X.[3] The present volume seeks to give an overview of the point this age-old debate between methodological rationalism and naturalism has reached, through interaction with contemporary psychology and wider reflection on what a proper philosophical naturalism should be. These developments promote and enable the development of an empirically grounded and psychologically informed metaphilosophy, a fresh metaphilosophical naturalism.

This introduction will provide background information about the evolution of methods in analytic philosophy, about relevant notions of 'intuition', and about the key projects of experimental philosophy, which are necessary to understand and contextualize current metaphilosophical debate. This debate is largely based on the assumption that, for better or worse, intuitions play a central role in analytic philosophy. As one elegant statement puts it:

> George Bealer does it. Roderick Chisholm does it a lot. Most philosophers do it openly and unapologetically, and the rest arguably do it too, although some of them would deny it. What they all do is appeal to intuitions in constructing, shaping, and refining their philosophical views.

> (Kornblith 1998, 129)

This 'centrality of intuition' assumption is (as Kornblith intimates) explicitly maintained more widely in metaphilosophical debate than in first-order philosophical discussion. Indeed, it has

recently attracted some very explicit criticism (Williamson 2007; Earlenbaugh and Molyneux 2009; Deutsch 2010; Cappelen 2012). Against such criticism, the first sections of this introduction will proceed from a nutshell history to explain why the 'centrality of intuition' assumption provides a potentially highly productive basis for methodological debate—despite the evident fact that philosophers, including mainstream analytic philosophers, do much more besides eliciting, marshalling, and weighing intuitions. We submit that the centrality assumption articulates a 'theme' that facilitates the development of competing 'models' (or paradigms, or exemplars) that have the potential to provide philosophy with distinctive methodologies beyond offering philosophers the licence to just follow their argumentative noses, as it were.

We shall now explain the relevant notions (Section 1) and recount how the centrality assumption became a philosophical 'theme' (Section 2, pp. 9–13). We will then explain in more detail different notions of 'intuition' and why intuitions matter (Section 3, pp. 13–20). Once we have thus defended the basic assumption that informs ongoing methodological debate and enables experimental philosophy, we shall give an overview of the main strands of this philosophical movement (Section 4, pp. 20–23), and outline how it catalyzes the emergence of a new *meta*philosophical naturalism and transforms the ongoing methodological debate between 'first-order' rationalism and naturalism (Section 5, pp. 23–27).

1 The methodological challenge, models, and themes

The key methodological challenge in any academic discipline—in any *Wissenschaft*—consists in the development of teachable methods and techniques that facilitate the achievement of consensual and correct solutions through replicable procedures of experiment, analysis, or argument. A caricature of our subject's development (cf. Austin 1956/1961, 232) may succinctly convey why this challenge has been particularly pressing in philosophy.

It is not entirely wrong to see philosophy as we now know it

as resulting from a process of continual specialization and fragmentation of intellectual inquiry, through which new intellectual communities emerged from a philosophical community that had initially concerned itself with all theoretical problems under the sun and developed wholly generic forms of argument that apply to any problem or topic. As more distinctive methodologies for dealing with specific kinds of problems matured, some new intellectual communities evolved into new disciplines; more recently, others led to the formation of distinctive new subdisciplines within philosophy, which often merged, or made close contact with, disciplines from beyond philosophy's previous remit, in response to 'area-specific' pressures and motivations (think, e.g., of the philosophy and history of science, mathematical logic, formal semantics, social epistemology, etc.). What problems philosophers work on hence depends on often contingent and still ongoing processes that have them shed, transform, and acquire questions and problems, without any general or overarching rationale. There is hence no reason to believe that the problems of contemporary philosophy are all of the same kind, or that claims about '*the* nature of philosophical problems' make much sense, or that it would be profitable to look for methods applicable to all such problems. At the same time, the repeated emigration of maturing specialist methodologies from philosophy has two consequences: First, the methodological challenge is more pressing for philosophy than for the disciplines that emerge from its midst precisely on the back and strength of a maturing methodology. Second, for many traditional areas of philosophy it takes the form of complementing wholly generic forms of argument that apply to any problem or topic with structured methods and techniques for dealing with specific kinds of problems.[4]

Many contemporary analytic philosophers would maintain that 'we do not have any such specific methods in philosophy, and do not need them.' There actually are such methods, though: Already the history of twentieth-century analytic philosophy provides us with a significant number of competing methods with the relevant degree of specificity. These methods are

typically embodied in *models* or 'paradigms', in something like the second and more fundamental of the two senses distinguished by Thomas Kuhn (Kuhn 1996, 175). These are achievements that are widely used, in teaching and research, as examples of one distinctive way of addressing one or more philosophically relevant tasks, which some philosophers hold up as representing models to be emulated in further work. Such models are provided by papers and, less often, books which are put on philosophy students' reading lists or are discussed in textbooks (which only acquired wider use in philosophy in the 1980s, though). They guide not only the students but also their teachers. They are models of what to do, rather than what to think. Hence a model of this kind can influence philosophers who disagree with the opinions of the author providing it, and the views it puts forward need not be popular, while most of the many texts that are influential at the level of content fail to acquire model status (e.g., because their authors just offer a compendium of arguments rather than follow any distinctive methodology, or because their methodologies are too subtle or sophisticated to be readily understood).[5] Such models influence, often implicitly, how philosophers address questions and problems, what kind of arguments and theories they develop, and what kinds of problems, arguments, and theories they are willing to take seriously.

The first such model to guide twentieth-century analytic philosophy, arguably, was Russell's 'On Denoting' (1905), which simultaneously served as a model of a certain approach (logical analysis) and provided (with Russell's theory of definite descriptions) a fresh model of what a specifically philosophical theory might look like, with Russell explicitly inviting counterproposals that would succeed as well as his own. Subsequently, several competing paradigms guided work in analytic philosophy, at almost any given point. Examples whose reception illustrates different ways in which models can guide philosophical work include: Carnap's 'Psychology in the Language of Physics' (1932/33), Moore's 'Proof of an External World' (1939), Ryle's *Concept of Mind* (1949), Austin's 'Plea for Excuses' (1956/57), Gettier's 'Is Justified True Belief

Knowledge?' (1963) in conjunction with classic responses to his challenge (such as Goldman 1976 and Lehrer and Paxton 1966), Davidson's 'Truth and Meaning' (1967), Kripke's *Naming and Necessity* (1972/1980), and a couple of papers on 'trolley problems' (Foot 1967) and the morality of abortion (Thomson 1971). The guidance through such models can be effective, even if typically implicit: We imitate those models that resonate with us, typically without explicitly telling ourselves (or our readers) that we will now try to do the same thing as some other author. These models influence our work to a larger extent and in more ways than we realize, and hence shape philosophical work more strongly than they shape philosophers' explicit self-conception.

While something like the second of Kuhn's notions of 'paradigm' applies to some philosophical works, we do not believe that full-fledged 'paradigms' in Kuhn's first sense of 'disciplinary matrix' guide philosophical work. This gap is (very) partially filled by slogans which shape philosophers' explicit self-conception and play a subtle but important role in the development and dissemination of fresh models. These slogans can helpfully be compared to musical themes: A *theme* is taken up again and again in a piece of music, by different instruments and in different variations; it lends coherence to the piece, which may contain so much more besides; and it serves as mnemonic and identifier for the piece, as it sticks in the mind most readily, is the first element of the piece to be remembered, and is therefore frequently used to pick out the piece and distinguish it from others. Until the 1970s, something along the line of the following slogan served the rich enterprise of analytic philosophy as a theme in such a way:

(A) We resolve philosophical problems through conceptual analysis that crucially includes linguistic analysis (but no psychological research).

Just as a musical theme is taken up and performed by different instruments, these words have been translated into intellectual deeds through various different kinds of logical, semantic, and

pragmatic analysis. Just as a recurring theme is played in different variations, these different kinds of analysis, and sometimes the same kind of analysis, have been employed for different philosophical ends: The sort of logical analysis pioneered by Russell, for instance, has been employed to answer some philosophical questions (e.g., Russell 1905) and show others meaningless (e.g., Carnap 1932/33); it has also been used to identify the structure of such different things as reality (e.g., Russell 1918) and statements about it (e.g., Hempel 1935; Quine 1953). More generally, analytic philosophers have tried to 'resolve' philosophical problems by defending pre-theoretical answers to questions, constructing philosophical theories, or exposing the problems as illusory—and each of these generic responses can take several different forms, many of which can be described as involving some sort of 'conceptual analysis'. Covering a potentially wide range of different, sometimes mutually incompatible, aims and methods, the slogan can be, and has been, interpreted in so many different ways that it does not offer a particularly informative description of the efforts of a philosophical community and gives little effective guidance to philosophical work.

Instead, such a theme provides an otherwise heterogeneous subject community with a—potentially misleading—sense of coherence, shared concerns, and direction, and serves to delineate the community, to decide who belongs to it and who does not. This decision, however, is not based on mere subsumption under the slogan. Rather, the slogan is used to sum up the most prominent strands of the multifarious enterprise of the group as a whole, which is then taken to consist of those philosophers who either pursue projects that can be subsumed under the slogan or meaningfully interact with proponents of these projects, if only by attacking them in a way that provokes responses from the attacked (think of Quine's propagation of naturalized epistemology). Hence the theme not merely fails to be particularly informative but, despite its wide coverage, fails to give an exhaustive description of the efforts of analytic philosophers at the time.

Despite these apparent shortcomings, the theme plays

important roles. For one, it serves to locate the community vis-à-vis other philosophical communities with which analytic philosophers do not meaningfully interact, and thus forges a group identity. Simultaneously, it serves to position the community thus forged vis-à-vis other disciplines. It does both by highlighting distinctive contribution(s) the subject makes that are worthwhile, distinct from those of other disciplines, and attainable by the means at philosophers' disposal. By promising to show them worthwhile, distinctive, and feasible, the theme legitimizes key strands of analytic philosophy.

Finally and crucially, the theme serves as a catalyst of methodological innovation: While it offers an umbrella for too many different methodologies to offer effective guidance for philosophical work on any specific question or problem, and for too few approaches to provide an exhaustive summary of the community's efforts, the theme does provide an avenue through which to launch fresh methodologies into the community: Fresh methodologies, explicitly advanced through programmatic explanations or demonstrated through new paradigms, can secure a serious hearing through subsumption under the theme, which simultaneously shows them legitimate and a part of the multifarious but still connected efforts of analytic philosophy.

In a nutshell, a theme serves a philosophical community not so much by giving an accurate description of its activities as by forging the community and providing a launch pad for fresh specific approaches. We now turn to the theme that guides current methodological debates in philosophy—which places intuitions at the centre of the subject.

2 A new theme

In the course of the 1970s, an increasing and eventually overwhelming number of anglophone philosophers came to disown the above theme (A), which seemed to constrict them to establishing facts at best about concepts, but not about the world (see, e.g., Armstrong 1977/1995, 175–177). These sentiments about exercises in conceptual analysis, as exemplified by efforts to analyse the concept of knowledge in

the wake of Gettier (1963), have been forcefully put:

> On the few occasions when I have taught the 'analysis of knowledge' literature to undergraduates, it has been painfully clear that most of my students had a hard time taking the project seriously … It was a source of ill-concealed amazement to these students that grown men and women would indulge in this exercise and think it important … For about as long as I can remember, I [too] have had deep, though largely inarticulate, misgivings about the project of analysing epistemic notions.
>
> (Stich 1990, 3)

Philosophers' disenchantment with conceptual analysis had a number of sources, quite apart from a lack of success in providing generally accepted analyses of specific concepts. One source was empirical research and related philosophical work that discredited the most prominent form of conceptual analysis, exemplified by the discussion of Gettier cases, namely the quest for definitions or sets of individually necessary and jointly sufficient application conditions: Important classes of concepts are not associated with such conditions; rather, their application is governed by prototypes (Rosch 1975), and the notion of concepts as definitions poses a range of empirical problems concerning their acquisition (Fodor 1975). Indeed, the relevant psychological research into the role of proto- and stereotypes in the application of classificatory terms strongly suggested that conceptual analysis needs to take into account empirical findings. So, what is philosophy to do, if it does not seek definitions through purely a priori reflection? Is it to pursue more comprehensive and naturalized forms of conceptual analysis, to revert to speculation, or to be a handmaiden to science (to mention just the most salient possibilities)?

The most prominent project of conceptual analysis to be (partially) naturalized and oriented away from the quest for classic definitions was the so-called 'Canberra Plan' (Lewis 1994; Jackson 1998). This project seeks to obtain the proper analysis of a concept *F* as part of a two-stage process. First, the

philosopher assembles and systematizes folk intuitions about F through 'armchair' reflection; second, the philosopher turns to science to discover what (if anything) all or most of these intuitive judgments are true of, so as to identify what F (say, colour, belief, or knowledge, etc.) is in the real world. The specific contributions of the philosopher thus consist in a systematic marshalling of intuitions and relating them to scientific findings.

Saul Kripke (1972/1980), Hilary Putnam (1975), and Tyler Burge (1979, 1986) had already proposed an externalism about content that promised to give intuitions a further role, outside conceptual analysis. For all three thinkers, what one understands by a concept or associates with the extension of it does not exhaust the concept's bearing on the truth of judgements involving it (there is more to a concept than what is in one's head or can be there). Such externalism is more or less plausible depending on the domain at issue (natural kinds as opposed to, say, numbers). It proceeds from such distinctively *modal* claims (explained below) as that (on the assumption that current chemical theory is correct) water necessarily is H_2O, independently of the occurrent properties we associate with samples of the compound. Thus, the truth of one's thoughts about water depends upon H_2O and its properties, even if one is chemically ignorant, like a denizen of the seventeenth century. Especially in the work of Kripke and many coming after him, intuitions and thought experiments acquire a new significance in the light of such an externalism: Whereas in the past intuitions were elicited to elucidate the content of concepts, they were now probed to determine *metaphysical* necessities. The philosopher, according to such 'modal rationalism', does not so much as analyse concepts independently of the way the world is, but uses intuitions to gauge and determine modal claims about how the world *must* be. This view provided an essential role for thought experiment and intuition independently of the increasingly unpopular enterprise of conceptual analysis.

A new theme then emerged from the development of such modal rationalism, from the advocacy of Canberra-style

analysis, and from reflection on the use of intuitions in philosophical thought experiments, more generally. The most philosophically prominent kinds of thought experiments have us consider the verbal description of certain scenarios and then make a judgment about it—which is then either used as premise in further argument (as in Jackson's thought experiment about the colour scientist Mary; Jackson 1982) or as evidence for or against a philosophical claim or theory (e.g., the standard understanding of Kripke's scenarios of deviant naming).[6] Indeed, even where other than modal or conceptual claims are at stake, philosophical theory construction frequently proceeds by working back and forth between intuitions elicited through different thought experiments and different kinds of background beliefs until 'reflective equilibrium' and a coherent set of judgments and beliefs is achieved (e.g., Rawls 1971; Thomson 1971; Foot 1967). Such an approach is sufficiently dominant for Bealer (1996, 4) to have labelled it analytic philosophy's 'standard justificatory procedure'. The common denominator of Canberra-style analysis, modal rationalism, and the use of thought experiments in philosophy more broadly seems to be captured by a new slogan:

(N)   Philosophers elicit, invoke, assess, and synthesize intuitions.

While some of the models initially subsumed under the previous slogan (A) can be equally happily subsumed under (N)—Gettier, etc.—others cannot (e.g., the works of Russell, Carnap, Moore, Ryle, Austin, and Davidson mentioned in Section 1, pp. 5–9). (N) is a new slogan, and no mere paraphrase of (A).

We will now show that this new slogan has all the characteristics of what we called a 'theme'. One upshot of this is that different models of intuition-based philosophizing play a key role in forging the community of current analytic philosophers, and open up fresh avenues for the introduction of fresh specific methodologies—so that intuitions and their study are absolutely central to philosophy, despite the evident fact that philosophers, including mainstream analytic philosophers, do much more

besides eliciting, marshalling, and weighing intuitions.

First, a themes' key terms are capable of, and subject to, multiple interpretations. This is true of all the key terms in (N). In current debates, philosophers have used over half a dozen different notions of 'intuitions', which have been taken to be

   (i) beliefs (Lewis 1983; van Inwagen 1997; Williamson 2007)
  (ii) judgments (Mercier and Sperber 2009; Ludwig 2007)
 (iii) inclinations to assent (Sosa 2007; Earlenbaugh and
          Molyneux 2009) or
  (iv) mental states entirely *sui generis* (Bealer 1998; Pust 2000)

with a particular kind of

 (a) phenomenology (Plantinga 1993; see also Parsons 1995)
 (b) justificatory status or justification (including 'conceptual
        accessibility') (Pust 2000; Bealer 2000; Goldman 2007;
        Sosa 2007; Ludwig 2010)
 (c) content (general or modal) (Bealer 2000, 3; Pust 2000;
        Sosa 2007) or
 (d) aetiology (Nagel 2012; Fischer 2014).

Methodological rationalists have us 'elicit' them through individual introspection or self-observation in thought experiments or the individually employed method of cases (critical review: Cappelen 2012), experimental philosophers through questionnaire-based surveys (overview: Alexander 2012), and pioneers of 'the sources project' (Nagel 2010, 2012; Fischer 2014) or 'iceberg epistemology' (Henderson and Horgan 2011) through psychological explanations. Intuitions are 'invoked' as evidence for or against philosophical claims and theories, about concepts and folk theories (Jackson 1998; Nahmias *et al.* 2005, 2006) and the phenomena these are about (Rawls 1971; Kripke 1972/1980; Shoemaker 1975), but also to motivate philosophical questions or raise philosophical puzzles, such as where philosophers' intuitions clash with background beliefs, so that the very possibility of what we take to be familiar facts comes to seem puzzling (Papineau 2009, and this volume,

Chapter 1; Fischer 2011). Also the 'assessment' of intuitions takes various different shapes. Philosophers seeking to achieve a narrow or wide 'reflective equilibrium' (Rawls 1974) assess intuitions according to their coherence with each other, with background beliefs, and with relevant theoretical considerations (Foot 1967; Thomson 1971; Rawls 1971). Others seek to assess intuitions empirically, by establishing their sensitivity or insensitivity to epistemologically otiose parameters like cultural and socio-economic background (e.g., Machery *et al.* 2004; Doris and Plakias 2008; Feltz and Cokely 2009), or to order and framing effects (e.g., Swain *et al.* 2008; Weinberg *et al.* 2012; Cushman and Schwitzgebel 2012). Yet other philosophers seek to derive assessments of our warrant for accepting intuitions from psychological explanations of why we have them, as and when we do (e.g., Nagel 2012; Fischer *et al.*, this volume, Chapter 12).

Second, while the precise extent to which philosophers elicit, appeal to, or discuss intuitions is unclear—on aetiological notions of intuition, for example, only the successful psychological explanation of certain judgments or beliefs will reveal whether they are intuitions—there can be little doubt that a non-negligible amount of philosophical work will fall outside the scope of the present slogan, on any reasonable interpretation of it: In such diverse areas as, for example, philosophical history and applied philosophy of science, intuitions are restricted to the kind of merely heuristic role they can have in any intellectual endeavour.

But, third, this slogan has highlighted features of philosophical work, such as reliance on thought experiments of the sort employed in the 'method of cases', or focus on paradoxes with intuitive premises, which are quite distinctive of philosophy in the analytic tradition (as opposed to other traditions) and of philosophy, *tout court* (as opposed to other subjects). Claims to distinctiveness on the latter front, however, may seem dubious in the light of apparently similar reliance on intuitions in other disciplines, first and foremost linguistics. Indeed, prominent philosophers of language (e.g., Devitt 2006) have regarded both philosophers and linguists as according privileged evidentiary

status to particular intuitions, namely their own, which are supposedly shaped by, and reflect, a special expertise (cf. Machery, this volume, Chapter 8). Linguists, however, do not regard whatever special expertise they might possess as relevant to the production of data for linguistic theory and, crucially, treat intuitive data just like any other data, with no especial epistemological privilege save one of convenience and freedom from known problems. (A linguist is not concerned with intuitions being true, only with them being robust.[7]) The new slogan (N) can therefore be used to draw attention to uses of intuitions that do set philosophers apart from other subject communities—including even linguistics.

Fourth, the new slogan has motivated different metaphilosophical positions and research programmes which seek to show or render different kinds of intuition-centred philosophizing feasible with means available to philosophers. Relevant efforts range from methodological rationalism which seeks to show intuition-based philosophizing possible with means available all along from the philosophical armchair (crucially including the exercise of conceptual competencies) to experimental philosophy which seeks to place new empirical methods at the disposal of philosophers who wish to elicit or assess relevant intuitions. Like a good theme, the slogan is thus highlighting distinctive contributions the subject makes that are worthwhile, distinct from those of other disciplines, and attainable by the means at philosophers' disposal.

Finally and crucially, this new theme has facilitated a new round of methodological innovation, led by experimental philosophy, which has already significantly increased the range of the means at philosophers' disposal. This ongoing introduction of empirical methods from the social sciences into philosophy may, in the long run, transform the subject at least as profoundly as the introduction of tools from formal logic transformed it a century ago.

3 Intuitions: what they are and why they matter

Just how philosophically productive the round of methodological

innovation facilitated by the new theme can be depends upon how much of, and how central, a role intuitions play in philosophical work. To assess their philosophical relevance, we first need to get clear on what is meant to be relevant. Different metaphilosophical positions and approaches have fashioned different notions of 'intuition' to meet their needs, resulting in the liberal usage canvassed above. We can order these notions on a spectrum, according to the extent to which they are informed by psychological research into intuitive judgment.

At the 'psychologically uninformed end' of the spectrum we find notions of 'rational intuition' that emerged from *modal rationalism*. This is the doctrine that intuition, or a priori reflection, can provide knowledge of such modal facts as whether a property necessarily (or only possibly) holds of an object or not (Stalnaker 2012). The doctrine is often credited to Kripke (1972/1980), but arguably goes back at least to Descartes. Kripke's immediate concern was to overthrow a particular philosophical theory of reference, known as 'descriptivism'. What endowed his criticism with wider significance, however, was that it involved separating the epistemological, the semantic, and the metaphysical. For Kripke, *a posteriori necessities* are not only coherent but readily witnessed in kind identity statements (such as 'Water is H_2O'); likewise, *a priori contingencies* are witnessed with descriptively fixed terms (e.g., '*Jack the Ripper* is whoever committed all *these* murders'). If necessity is *not* coeval with analytic truth, and so can be a posteriori (as in 'Water is H_2O'), then one needs a warrant for any claim of necessity beyond appeal merely to one's competence with the relevant words.[8] Accordingly, modal rationalism suggests that the warrant for the present arguments is provided not so much by semantic competence, but through modal insight, as it were, in terms of conceivability. There then appears to be a ready way of deciding on conceivability, namely, a thought experiment to show that a property could hold without contradiction.[9]

Modal rationalism is restricted to questions of modality: Kripke-style argument does not support the conclusion that

intuitions generally possess or afford a significant warrant. According to the metaphilosophical position known as *(modest) methodological rationalism* (most extensively developed by Bealer 1996, 2000), philosophers should content themselves with the sort of general and necessary claims that might be established through such intuitive insights, which can be plausibly traced to the exercise of broadly construed conceptual competencies: Philosophers should seek general and necessary answers to questions about the nature of things (the mind, perception, truth, causation, etc.), by constructing theories that achieve a reflective equilibrium between 'rational intuitions' that provide a priori justification for those necessary truths.

This led methodological rationalists to posit intuitions which can be reliably identified from the armchair and afford a priori justification for precisely the kind of general modal claims the rationalists are interested in. They typically did so without considering any psychological literature. As taken up more widely, the resulting *psychologically uninformed rationalist notion* requires that intuitions possess

(a) a distinctive phenomenology (Plantinga 1993),
(b1) a 'default justificatory status': they afford justification without requiring any themselves (Bealer 1996; Pust 2000), and
(b2) 'conceptual accessibility': their truth can be recognized through exercise of conceptual competencies alone (Bealer 2000; Goldman 2007; Sosa 2007; Ludwig 2010).

This notion is taken for granted in the recent objections to the centrality of intuitions assumption we mentioned at the outset. The most forceful of these is due to Herman Cappelen, who conducted a series of case-studies on supposed paradigm cases of intuition-based philosophizing and argued that, as a matter of empirical fact, analytic philosophers do not rely on judgments possessing *any* of the three properties listed, as evidence for their theories (Cappelen 2012, 111–187).[10] The best response to this forceful objection, we suggest, is to regard it as an open question whether *such* 'rational intuitions' even

exist, let alone play the central role in philosophy that the new theme (N) accords intuitions. We should, we submit, shift attention to psychologically informed notions and intuitions whose existence can be empirically demonstrated, and regard it as an empirical question whether such intuitions have a characteristic phenomenology or are conceptually accessible.

Contemporary cognitive psychology uses an aetiological notion of 'intuition'—(d) in the list on page 11, above.[11] In philosophy, this notion is used by approaches which fall within the 'submarine part' of 'iceberg epistemology' (Henderson and Horgan 2011). They seek to derive epistemological assessments of intuitions from their psychological explanations, in particular from explanations that trace intuitions back to automatic cognitive processes that take place 'below the waterline' of conscious awareness. Their *psychologically informed aetiological notion* explicates intuitions as judgments which are

(1) based on largely automatic inferences (Kahneman and Frederick 2005, 268; Sloman 1996; see also Evans 2010, 314), namely on largely automatic cognitive processes which duplicate rule-governed inferences,[12] and
(2) accompanied by 'feelings of rightness' (Thompson *et al*. 2011) (i.e., which immediately strike the thinker as plausible, regardless of whether or not she accepts those judgments upon further reflection).

The relevant notion of automaticity is gradual, rather than dichotomous: Processes are more or less automatic depending upon the extent to which they are effortless, unconscious, non-intentional, and autonomous. These four properties are all individually gradable and operationally defined (Bargh 1994; Moors and De Houwer 2006; see also Dijksterhuis 2010): A process possesses, for example, the key property of being effortless to the extent to which it requires no attention or other limited cognitive resource, so that performance is not impaired by multitasking (keeping in mind long numbers or complex dot patterns, etc.).[13] The attendant 'feelings of rightness' are

typically identified through spontaneous assessments of subjective confidence, where subjects are asked to indicate on a Likert scale whether when making the relevant judgments they 'felt guessing', 'fairly certain' or 'certain I'm right'.

That a philosopher's judgment is an intuition in this sense can only be established through a successful, experimentally confirmed, psychological explanation that traces it back to automatic cognitive processes. While pertinent explanations of philosophically relevant judgments begin to get developed (e.g., Nichols and Knobe 2007; Fiala *et al.* 2011; Nagel 2012; Fischer *et al.*, this volume, Chapter 12), it is still too early to make confident assertions about the extent to which philosophers rely on intuitions, in psychologists' aetiological sense. There are, however, some quite compelling prima facie reasons to believe that such intuitions do play a key role in contemporary analytic philosophy. We will now review those reasons, to show that, properly interpreted, the new theme (N) is descriptively reasonably accurate (to the extent to which themes are, see Section 1, pp. 5–9) and potentially productive.

While methodological rationalists have not been able to offer reasonably informative descriptions of any 'distinctive phenomenology' that would allow us to reliably identify trustworthy judgments, the judgments we make about the scenarios we consider in applying the method of cases do have some characteristic phenomenological traits: We tend to make them spontaneously, i.e., (i) swiftly and (ii) effortlessly, (iii) with significant initial confidence, and (iv) they have a way of seeming plausible to us even once we have subsequently decided they are wrong (as and when we do), a bit in the way in which the lines in a Müller–Lyer diagram continue to look a different length even once we know they are the same.[14] Second, while normative foundationalist claims and attributions of 'default justificatory status' may be motivated primarily by the rationalist agenda, many philosophical case-judgments (e.g., about Gettier or trolley cases) are, as a brute matter of fact, typically (v) accepted immediately and (vi) without asking for further justification or offering any argument. Rather, philosophers are required to honour them; i.e., the requirement

that they motivate or justify the given judgment is imposed on any principles subsequently adduced to rationalize it.[15] Indeed, in many such cases (again, think of Gettier or trolley cases) the subsequent formulation of principled grounds licencing the initial judgments proves difficult. Here, the judgments initially elicited by the vignette were arguably made (vii) without awareness of reasons the philosophers at issue would be willing to endorse.

Intuitions in the aetiological sense have all these features (i–vii)—and contemporary psychology can explain why. These explanations make use of a framework widely adopted also in experimental philosophy: 'Dual-process theories' (reviews: Evans 2008; Evans and Stanovich 2013) distinguish two kinds of cognitive processes: Rapid automatic ('type-1') processes place few demands on the resources of working memory, have the above-listed process-properties to a high degree, and involve the execution of several steps in parallel. Slower controlled ('type-2') processes (like mental arithmetic) rely on working memory, are more effortful (you get your sums wrong when distracted), conscious, intentional and controlled, and go on in accordance with rules, in a serial fashion, one step at a time. Automatic processes produce judgments and decisions which we accept as a default; only in response to particular cues do we engage in conscious reflection which may result in correction or other modification of the initial rapid response.[16] Dual-process theory can explain, for a start, why philosophers feel torn between characterizing spontaneous judgments about vignettes in thought experiments as judgments (e.g., Mercier and Sperber 2009; Ludwig 2007) and as inclinations to judge or assent (e.g., Sosa 2007; Earlenbaugh and Molyneux 2009): Where judgments issuing from automatic cognitive processes get immediately accepted, there is no problem with describing them as, well, judgments. Where pertinent cues (see below) prompt swift conscious correction or modification of initial responses,[17] it seems more appropriate to speak of 'inclinations to judge'. In either case, we are dealing with intuitions, in the psychologists' sense.

Against common prejudice, however, conscious reflection is

not necessarily superior to prior automatic processing: Both of the major research programmes on intuition in cognitive psychology, namely the 'heuristics and biases programme' (Tversky and Kahneman 1974; Kahneman and Frederick 2005; Kahneman 2011) and the 'adaptive behaviour and cognition programme' (aka 'fast and frugal heuristics'—Gigerenzer *et al*. 1999; Gigerenzer 2008) converge on the finding that the automatic processes studied issue in reasonably accurate judgments under most ordinary circumstances, and helpfully deliver such judgments even under circumstances where the rules we consciously master require information or resources we do not have, and thus leave us in the lurch, while post-intuitive reflection need not correct intuitive judgments in the cases in which they do go wrong, but may as well result in confabulation and the formulation of dubious *ex post* justifications (an alternative neglected in psychologically informed philosophical debates) (Shynkaruk and Thompson 2006; Stanovich 2009).

By the above definition, intuitions in the aetiological sense have the characteristic phenomenology of philosophical case-judgments: being (1) generated by automatic processes, they are (i) rapid and (ii) effortless, while (2) accompanied by (iii) high levels of subjective confidence or 'feelings or rightness'. Psychological work on metacognitive cues within the dual-process framework lets us understand how and why these features hang together: According to the now dominant 'experience-based approach' to metacognitive judgments (review: Koriat 2007), the subjective confidence attaching to spontaneous responses does not result from deliberate reflection on their content or further information retrieved from memory. Rather, it results directly from features of the process that generate the judgment. The most important of these is the ease with which this conclusion of an automatic inference comes to mind (known as 'answer fluency') (Thompson *et al*. 2011, 111; see also Simmons and Nelson 2006). This ease is operationalized in terms of response time and effortlessness. A subject's confidence in a judgment increases with the speed with which she arrives at it (Kelley and Lindsay 1993; Robinson

et al. 1997; Thompson *et al.* 2011, 2013) as well as with the subjective impression of effortlessness (Alter *et al.* 2007). Since the conclusion of the automatic inference will continue to come to mind readily, even if, as, and when we have rejected it upon subsequent reflection, the 'feeling of rightness' will persist and the judgment will (iv) continue to strike us as intuitively plausible.[18] The characteristic phenomenology of philosophical case-judgments therefore follows from satisfying the aetiological definition of 'intuitions' and forms a coherent whole.

The subjective confidence engendered by fluency serves, in turn, as a cue that determines whether a subject accepts an initial intuitive judgment without further ado or engages in conscious reflection: The more confident we feel about it, the more likely we are to accept a spontaneous judgment; the less confident we feel, the more time we spend scrutinizing it (Thompson *et al.* 2011, 2013). Hence highly intuitive judgments are particularly likely to be (v) accepted immediately, without engaging in effortful reflection and explicit argument. Subjective confidence also serves as a consensus cue: The more confident we are about a judgment, the more likely our response is to be shared by others, and the less controversial we—rightly—take our judgment to be (Koriat 2008, 2012). The less controversial a judgment is, the less pragmatic need there is for supporting argument, and this makes us even more inclined to (vi) assert it without such argument.[19] Since these spontaneous judgments are engendered by automatic processes into which we have little, if any, insight, we cannot subsequently report any reasons we had in making our judgments and (vii) may find it difficult to construct a justification we find acceptable. The suggestion that they are intuitions in the aetiological sense of the term would seem to offer the best available explanation of the salient traits of philosophical case-judgments we indicated (i–vii).[20]

Judgments which appear to have these traits are not only made about cases philosophers consider in the context of thought experiments or in applying the method of cases for purposes of conceptual analysis, etc. They also figure as

premises in philosophical paradoxes that motivate characteristically philosophical questions of the form

How is it (possible) that *p* (given that *q*)?

where *q* is the last step of an apparently sound argument and *p* is generally accepted as a familiar fact, but which appears to be inconsistent with *q* (Fischer 2011; cf. Papineau 2014 and this volume, Chapter 1). These questions articulate the kind of puzzlement in the face of the familiar that Plato famously regarded as the beginning of all philosophizing (Fischer 2011, 206–210). Examples include sceptical paradoxes which have us wonder how it is possible that we acquire knowledge through our senses (Greco 2007), or about others' beliefs and desires (Avramides 2001), or about the past (Ayer 1956), classical paradoxes about mental causation or free will, which have us wonder how it is possible for our beliefs and desires to make any difference to our bodily movements (Maslen *et al*. 2009), or how we can possibly be morally responsible for anything (Kane 2011), as well as different paradoxes that jointly make up the problem of perception: 'How is it possible for us to perceive physical objects and public events, given that all we are directly aware of are subjective perceptions?' (Smith 2002, discussed by Fischer *et al.*, this volume, Chapter 12). Intuitive judgments may be philosophically relevant either as the evidential basis of philosophical theories or as the root of philosophical problems or in yet further ways.

Which judgments are relevant to philosophy in one or more of these ways can be established through case studies on paradigmatic developments of philosophical problems, arguments, and theories, in influential philosophical texts (such as those undertaken by Fischer 2011 and Cappelen 2012). The precise extent to which these judgments actually possess the traits that allow us to characterize them as intuitions in the aetiological sense can be rigorously determined only through psychological experiments which measure the speed with which they are made, the extent to which they are affected by multitasking, etc. Such research has already been done for

various judgments consistent with well-researched heuristics (e.g., De Neys 2006) but remains a *desideratum* for the many different intuitions at issue in philosophical debates.

The present considerations do, however, support the hypothesis that intuitions, aetiologically conceived, play a key role in philosophy, regardless of whether or not philosophers manage to have 'rational intuitions' of the kind that proved elusive (Cappelen 2012).

4 Experimental philosophy

Aetiological notions of intuitions allow us to draw on extant and ongoing psychological research. Over the last fifteen years or so, an increasing number of philosophers have come to draw on such research and imported its methods into several branches of philosophy. While methodological naturalism has long enjoined philosophers of mind to take into account findings from psychology and the other cognitive sciences, experimental philosophers started to employ also the methods of psychology and other social sciences, and did so in pursuit of questions from epistemology, metaphysics, moral philosophy, and the philosophies of action and language, in addition to the philosophy of mind.[21] Their efforts can be summed up by a subtheme of the new theme (N):[22]

(E) Experimental philosophers employ findings and methods from the social sciences, crucially including psychology, to elicit, explain, and assess philosophically relevant intuitions.

Work falling under this theme is diverse: It employs different means, in pursuit of different ends.

The freshly imported means employed fall into two broad categories: questionnaire-based *surveys* that establish correlations and *experiments* that involve the active manipulation of a relevant variable.[23] In a typical experimental-philosophy survey, participants are asked about their ethnic or educational background, etc., are given a vignette that describes a philosophically interesting scenario (say, a Gettier

or trolley case), and are asked which of different judgments about this scenario (*X* knows/does not know, *X* is morally good/neutral/bad) they spontaneously deem correct (e.g., Weinberg *et al.* 2001; Cushman and Schwitzgebel 2012). This may reveal that members of different ethnic groups, etc., tend to make different judgments about these cases. In a simple experiment, one might randomly assign participants to different groups and present them with the same cases in a different order (e.g., Swain *et al.* 2008). This may confirm the hypothesis that people's judgment about a particular case is affected by what cases they have considered before. While a first generation of contributions to experimental philosophy employed mainly surveys and such simple experiments, a nascent second generation is drawing on findings from more sophisticated experiments undertaken by cognitive and social psychologists (e.g., Nagel 2012) and begins to employ the relevant experimental paradigms themselves (see Fischer *et al.*, this volume, Chapter 12).

These means are deployed towards three different ends: They serve (1) to establish truths about folk concepts and folk theories guiding the application of concepts, (2) to assess the evidentiary value of philosophically relevant intuitions, and (3) to establish substantive facts, namely, about human psychology. Some examples may illustrate the practice of experimental philosophy as (1) conceptual analysis (in a liberal sense), (2) epistemological assessment, and (3) psychology:

(1) The concept project: In the analysis of the concept of knowledge, surveys have been used to test different contextualist analyses of knowledge, namely, to determine whether non-philosophers' attributions of knowledge are sensitive to explicit mention of specific error possibilities (Buckwalter 2010) or to how much is at stake for the protagonists (Feltz and Zarpentine 2010). In the philosophy of action, surveys have been used to elucidate the folk concept of intentional action (Alexander 2012, 59–69) and led to the surprising finding that attributions of intentions to act are linked to moral assessments of the action in question

(Knobe 2003). In moral philosophy, surveys have been used to expose key components of folk theories of responsibility, such as whether people regard determinism as compatible or incompatible with free will and moral responsibility (Nahmias *et al.* 2004, 2005, 2006).

(2)   The warrant project*:* Both surveys and experiments have been used to assess intuitions' *evidentiary value*, i.e., whether the mere fact that given thinkers have a particular intuition, as and when they have it, speaks for its truth. Some philosophers seek positive evaluations, to defend the use of the targeted intuitions as evidence for philosophical theories (e.g., Nagel 2012). Others seek negative evaluations, either to attack this use (e.g., Swain *et al.* 2008), or to resolve philosophical paradoxes by showing that we lack warrant for accepting some of their intuitive premises in the common absence of further argument (e.g., Fischer *et al.*, this volume, Chapter 12).

(2.1)   First generation: Pertinent questionnaire-based surveys have been guided by the basic idea that intuitions lack such value (a) when they vary between different groups that do not differ in epistemic position or credentials, and (b) when they are sensitive to factors that have nothing to do with the content of the judgment and therefore ought not to affect the judgment (reviews: Alexander and Weinberg 2007; Alexander 2012). In this vein, experimental philosophers have examined whether intuitive attributions of reference or knowledge vary between Western and East Asian students (Machery *et al.* 2004; Weinberg *et al.* 2001) and whether people's spontaneous moral judgments about cases change when the cases are presented in different order or formulated in different, but equivalent terms (Weinberg *et al.* 2012; Petrinovich and O'Neill 1996). The basic idea informing this approach, in particular part (a), has attracted significant criticism (including Wright 2010; Shieber 2010; Pinillos *et al.* 2011; Williamson 2011).

(2.2) Second generation: A different approach that does not rely on this basic idea has been characterized as 'the sources project' (Pust 2012) or 'cognitive epistemology' (Fischer 2014): It synthesizes and extends results from psychological experiments to formulate psychological explanations of philosophically relevant intuitions that help us assess their evidentiary value—and our warrant for accepting them in the absence of further argument. One key idea here is to show that the intuitions explained (e.g., Gettier-intuitions) issue from a cognitive process which is generally reliable, even if it occasionally leads to cognitive illusions. The processes examined include mind-reading (Nagel 2012), non-intentional analogical inference (Fischer 2014, 2015), and stereotype-driven amplification (Fischer *et al.*, this volume, Chapter 12).[24] Such explanations yield a positive assessment of intuitions generated under normal or propitious conditions, and a negative assessment of intuitions generated under the special conditions that lead to cognitive illusions.[25]

(3) The psychology project*:* Some problems traditionally discussed by philosophers can readily be treated as problems of psychology. One example is the 'descriptive (as opposed to normative or sceptical) problem of other minds': What makes us think that others enjoy conscious mental states like beliefs and desires? One group of experimental philosophers proposed, against current abductive orthodoxy, that simple behavioural features (including motion trajectories and contingent interaction) trigger automatic attributions of agency which prime, or dispose subjects towards, attributions of conscious mental states. Since such automatic processes would often trigger attributions to inappropriate 'agents' (such as moving dots on a screen), requiring effortful suppression, the proposal was tested through response-time measurements for appropriate and inappropriate attributions (Arico *et al.* 2011).[26]

Experimental philosophy is often regarded as the epitome of methodological naturalism. In fact, however, the first two of the three projects indicated can be pursued with the aim of justifying methodological rationalism no less than to vindicate methodological naturalism. When practising experimental philosophy as conceptual analysis, one might conduct surveys of the folk in the hope of showing that their intuitions are as similar to those of mainstream analytic philosophers as competent speakers' intuitions about the grammaticality of sentences are to the analogous intuitions of expert linguists, so that philosophers do not need to consider intuitions other than their own to gain insight into folk concepts—or perhaps even the phenomena these concepts stand for (see Shieber 2012). Other philosophers, by contrast, welcome findings that actual folk theories diverge from those philosophers postulated on the basis of their own intuitions (Nahmias *et al.* 2004, 2005, 2006), and infer that armchair intuiting fails to warrant claims about folk concepts and theories.

Similarly, one can engage in epistemological assessment of 'armchair' intuitions both with the aim of showing that we may not trust such intuitions and with a view to showing that we have warrant to rely on—particular—armchair intuitions (take, e.g., the several contributions critically discussed by Kornblith, this volume, Chapter 6). Indeed, the official rationale of the 'restrictionist project' that employs surveys to uncover otiose sensitivities (2.1 above) is precisely to delineate the proper domain for reliance on armchair intuition (e.g., Weinberg 2007). The same goes for psychological explanations that explain a class of intuitions as resulting from cognitive processes that are generally reliable but engender cognitive illusions under specific, predictable circumstances (2.2 above): They can be developed with the aim of vindicating reliance on the intuitions explained (e.g., Nagel 2010, 2012) or in order to expose them as cognitive illusions we have no warrant to accept (e.g., Fischer 2014, 2015; Fischer *et al.*, this volume, Chapter 12). To sum up, experimental philosophy is not a party to the dispute between methodological rationalism and naturalism, but offers a new framework for settling it.

5 Metaphilosophical naturalism

What makes this new framework so exciting is that it provides both philosophical and metaphilosophical discussion with a fresh set of specific methods and distinctive approaches beyond 'just following our argumentative noses': It imports specific scientific methods (namely surveys, behavioural experiments, and their statistical analysis) which are reasonably well understood, and puts these methods to distinctively philosophical uses (such as conceptual analysis or epistemological assessment, for the purposes of philosophical theory construction or problem resolution). In particular, it provides us with the framework of an empirically grounded metaphilosophy that facilitates a differentiated and piecemeal approach: It does not assume that all philosophical questions and problems are of one and the same kind, but allows, for example, that there are some philosophical questions which are appropriately answered by theories and some that are appropriately 'dissolved' through diagnostic resolution of paradoxes that motivate their formulation—to name but two of several possibilities. And it allows us to delineate precisely where and when philosophers may rely on intuition and armchair methods, in their various endeavours—rather than issuing any wholesale bans or licenses.

This fresh framework has facilitated novel debate between methodological rationalists and naturalists, at three levels: debate within the framework, debate about the framework, and debate motivated by the framework. Debate within the framework is advanced mainly through different contributions to the project of epistemological intuition assessment (2.1 and 2.2 above). Debate about the framework focuses on the questions of whether the kind of surveys (2.1) or psychological explanations (2.2) conducted or constructed by experimental philosophers examine the right kind of intuitions and establish the right sort of conclusion, to have the kind of philosophical relevance to which they aspire. This debate thus revolves around a '*twofold relevance challenge*'.

First, several opponents of experimental philosophy have

raised '*the target issue*' (as we might call it): They have questioned whether those surveys and explanations examine the kind of intuitions that matter in philosophy. The 'reflection' and 'expertise objections' are particularly fundamental in applying to both approaches: According to the former, only the intuitions of careful and reflective thinkers are relevant in philosophy (e.g., Kauppinen 2007, 97); according to the latter, only the intuitions of philosophical experts matter (e.g., Ludwig 2007). In either case, surveys of hasty folk intuitions are as irrelevant as psychological explanations based on experiments on mere psychology undergraduates (the usual members of participant pools). In response, experimental philosophers have drawn on concepts and findings from cognitive psychology to assess the merit of these objections. For example, Machery (this volume, Chapter 8) reviews results from psychological expertise research and identifies specific and substantive reasons to believe that philosophers may be prone to an 'illusion of expertise' (see also Clarke 2013). Similarly, Weinberg and colleagues (2012) operationalized the notion of 'careful and reflective thinkers' through a common psychological measure known as the Need for Cognition Scale and found that the intuitions of 'reflective' subjects who scored high on this scale were subject to order effects to roughly the same extent (if in a surprisingly different way!) as those of less reflective subjects. This illustrates a key strength of experimental philosophy: the ability to bring empirical methods and findings to bear on many metaphilosophical questions, which may thus get actually settled, as pertinent scientific evidence gradually accumulates.[27]

Second, experimental philosophers themselves have raised the question of how their surveys, etc., can establish conclusions of the kind they ultimately seek. A manifesto of the movement succinctly sums up this *basic methodological puzzle about experimental philosophy*:

> In a typical experimental philosophy paper, the evidence being gathered is about the percentages of people who hold various sorts of intuitions, but the theories under

discussion are not about people's intuitions but about substantive philosophical questions in epistemology, metaphysics, or ethics. It may appear, at least on first glance, that there must be some sleight of hand involved here. How on earth could information about the statistical distribution of intuitions ever give us reason to accept or reject a particular philosophical view?

(Knobe and Nichols 2008, 6)

This worry is particularly live for the 'warrant project' of epistemological assessment—(2) above—which seeks to establish normative conclusions about what right or warrant thinkers have to accept certain philosophically relevant intuitions. Many experimentalists are ultimately interested, not in intuitions about *X*, but in the rights and wrongs of particular philosophical views about *X*. But how could findings about the statistical distribution of intuitions ever license conclusions about what warrant we have to accept, first, certain philosophically relevant intuitions and, second, particular philosophical views? If this strikes you as a burning question, read on: It is the guiding question of the second part of this volume (see the following synopsis).

Finally, the interaction with cognitive and social psychology catalyzed by experimental philosophy has motivated metaphilosophical debate beyond the framework provided by the three strands (1–3 above) of card-carrying experimental philosophy. In particular the sources project has been actively pursued by philosophers who are no card-carrying members of that movement, and some philosophers have pursued efforts in keeping with the spirit of that project but of wider scope. For example, Georges Rey (1998), Alvin Goldman (1999), and Louise Antony (2004) inquired, more generally, into the empirical, psychological conditions of the possibility of a priori judgment and knowledge.

Crucially, attention to scientific psychology in metaphilosophical debate led to a wider appreciation of the relevance of empirical studies beyond psychology, for metaphilosophical questions. Thus, since experimental

philosophers turned to empirical methods to assess the evidentiary value of intuitive judgments, friends and foes of the movement have started to undertake case studies on influential philosophical texts to find out what use or uses philosophers actually make of intuitions, at several stages of their work, from the formulation of philosophical problems to the construction and assessment of philosophical theories (e.g., Fischer 2011; Cappelen 2012). These studies employ familiar techniques of textual exegesis and logical reconstruction, which can be deployed from the armchair; but these armchair tools are deployed to establish empirical claims (just like the tools of the historian's trade, an equally empirical enterprise), namely claims about how particular philosophers, people in flesh and blood, actually reason and work. Experimental philosophy has thus promoted, through attraction or repulsion, a wider *metaphilosophical naturalism* that asks that we proceed from scientific and, more generally, empirical findings when addressing such metaphilosophical questions as what kind of problems philosophers address and how they come to conceptualize them, what kind of theories they seek to construct, what methods they do and should employ, what kinds of argument or evidence are relevant for which of their concerns, and what evidence they do and may rely on under which circumstances.

The interaction with cognitive and social psychology catalyzed by experimental philosophy has transformed both methodological rationalism and naturalism. On the rationalist side, it has led to the development of psychologically informed defences of forms of rationalism that are more aware of several empirical enabling conditions of a priori judgment and knowledge. At the same time, it has given methodological naturalism a new twist: Where traditional 'first-order naturalists' seek to address philosophical problems about X by taking into account scientific findings about X, experimental philosophers proceed by taking into account—and generating—scientific findings about the way we (philosophers) think about X.

This implies some commitment to what we have just called 'metaphilosophical naturalism'. But *such* naturalism requires no

Descartes, R. 1637/1993. *Discourse on Method*. In *Discourse on Method and Mediations of First Philosophy*, trans. D. A. Cress, 3rd ed. (pp. 1–44). Indianapolis: Hackett.

Deutsch, M. 2010. Intuitions, counterexamples, and experimental philosophy. *Review of Philosophy and Psychology*, 1: 447–460.

Devitt, M. 2006. *Ignorance of Language*. Oxford: Oxford University Press.

Dijksterhuis, A. 2010. Automaticity and the unconscious. In S. T. Fiske, D. T. Gilbert and G. Lindzey (eds.), *Handbook of Social Psychology*, 5th ed. (pp. 228–267). Hoboken: Wiley.

Doris, J. and Plakias, A. 2008. How to argue about disagreement: evaluative diversity and moral realism. In W. Sinnott-Armstrong (ed.), *Moral Psychology*, vol. 2: *The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, MA: MIT Press.

Earlenbaugh, J. and Molyneux, B. 2009. Intuitions are inclinations to believe. *Philosophical Studies*, 145: 89–109.

Evans, J. S. B. T. 2008. Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology*, 59: 255–278.

Evans, J. S. B. T. 2010. Intuition and reasoning: a dual-process perspective. *Psychological Inquiry*, 21: 313–326.

Evans, J. S. B. T. and Stanovich, K. E. 2013. Dual-process theories of higher cognition: advancing the debate. *Perspectives on Psychological Science*, 8: 223–241.

Feltz, A. and Cokely, E. T. 2009. Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Consciousness and Cognition*, 18: 342–350.

Feltz, A. and Zarpentine, C. 2010. Do you know more when it matters less? *Philosophical Psychology*, 23: 683–706.

Fiala, B., Arico, A. and Nichols, S. 2011. The psychological origins of dualism. In E. Slingerland and M. Collard (eds.), *Creating Consilience: Integrating Science and the Humanities* (pp. 88–110), New York: Oxford University Press.

Field, A. and Hole, G. 2003. *How to Design and Report Experiments*. London: Sage.

Fischer, E. 2011. *Philosophical Delusion and Its Therapy*. New York: Routledge.

Fischer, E. 2014. Philosophical intuitions, heuristics, and metaphors. *Synthese*, 191: 569–606.

Fischer, E. 2015. Mind the metaphor! A systematic fallacy in analogical reasoning. *Analysis*, 75: 67–77.

Fodor, J. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Foot, P. 1967. The problem of abortion and the doctrine of double

Kahneman, D. 2011. *Thinking Fast and Slow*. London: Allen Lane.

Kahneman, D. and Frederick, S. 2005. A model of heuristic judgment. In K. J. Holyoak and R. Morrison (eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 67–293). Cambridge: Cambridge University Press.

Kane, R. (ed.) 2011. *The Oxford Handbook of Free Will*, 2nd ed. New York: Oxford University Press.

Kant, I. 1787/2003. *Critique of Pure Reason*, 2nd ed., trans. N. Kemp Smith. Basingstoke: Palgrave.

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

Kelley, C. M. and Lindsay, D. S. 1993. Remembering mistaken for knowing: ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, 32: 1–24.

Knobe, J. 2003. Intentional action and side effects in ordinary language. *Analysis*, 63: 190–193.

Knobe, J. and Nichols, S. 2008. An experimental philosophy manifesto. In J. Knobe and S. Nichols (eds.), *Experimental Philosophy* (pp. 3–14). Oxford: Oxford University Press.

Knobe, J. and Prinz, J. 2008. Intuitions about consciousness: experimental studies. *Phenomenology and the Cognitive Sciences*, 7: 67–83.

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

Koriat, A. 2007. Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch and E. Thompson (eds.), *The Cambridge Handbook of Consciousness* (pp. 289–326). Cambridge: Cambridge University Press.

Koriat, A. 2008. Subjective confidence in one's answers: the consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34: 945–959.

Koriat, A. 2012. The self-consistency model of subjective confidence. *Psychological Review*, 119: 80–113.

Kornblith, H. 1994. Naturalism: both metaphysical and epistemological. *Midwest Studies in Philosophy*, 19: 39–52.

Kornblith, H. 1998. The role of intuitions in philosophical inquiry. In M. DePaul and W. Ramsay (eds.), *Rethinking Intuition* (pp. 219–141). Lanham, MD: Rowman & Littlefield.

Kripke, S. 1972/1980. *Naming and Necessity*. Oxford: Blackwell.

Kuhn, T. S. 1996. *The Structure of Scientific Revolutions*, 3rd ed. Chicago: University of Chicago Press.

Kutas, M. and Federmeier, K. T. 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends in*