

Introduction: New Perspectives on Philosophical Thought Experiments

Adriano Angelucci¹ · Margherita Arcangeli²

Published online: 18 August 2018 © Springer Nature B.V. 2018

"If philosophy misconceives what it is doing"—as an influential contemporary philosopher has wisely observed—"it is likely to do it worst" (Williamson 2007, Preface ix, emphasis added). The same would arguably apply if philosophers were to misconceive how they do what they are doing. Over the last couple of decades, many philosophers working in the analytic tradition have felt the pressing need to address anew questions concerning the nature and limits of philosophical inquiry. What exactly is analytic philosophy primarily concerned with? What are its proper methods? Can the kind of questions typically asked by analytic philosophers be satisfactorily answered "from the armchair"? Are the methods of the social and natural sciences fit to solve—or at least to illuminate—traditional philosophical problems? If so, how? And if not, why? Similar long-standing foundational questions have recently begun to insistently resurface, and to take centre stage, in many philosophical debates, thereby opening a lively new season of meta-theoretical reflection (cf., e.g., Jackson 1998; Williamson 2007; Cappelen 2012; Deutsch 2015; Machery 2017). This powerful surge of interest in metaphilosophy spawned a steadily growing body of literature devoted to the careful study of recent philosophical practice, and analytic philosophy has entered a phase of systematic reassessment of what it was previously, and often uncritically, taken to be its standard methodology (cf. DePaul and Ramsey 1998; Haug 2014; Fischer and Collins 2015; Cappelen et al. 2016; Nado 2016; D'oro and Overgaard 2017). The main goal of the present Issue is that of contributing to this ongoing process by focusing on the pervasive use of *thought experiments* (henceforth TEs)—often referred to as the *method of cases*—in analytic philosophy.

It can hardly be denied that TEs constitute a fundamental item within the bag of tools of analytic philosophers. While their use is certainly not new to Western thought (cf., e.g., Horowitz and Massey 1991; Stich and Tobia 2016), philosophical theorizing in the analytic tradition has been characterized by a massive and confident reliance on the method of cases. During the second half of the twentieth Century, in particular, the critical scrutiny of hypothetical scenarios usually, though not always, conceived as part of the traditional philosophical project of conceptual analysis—began to play an increasingly prominent role in the development and assessment of many philosophical views. The agreement of a philosophical analysis, account, or theory with our judgments about which facts would hold in a given casei.e. the pursuit and achievement of a reflective equilibrium between the two—was commonly regarded both as a reliable indicator of truth, and as a highly desirable theoretical virtue. Philippa Foot's trolleys (Foot 1967), Edmund Gettier's cases (Gettier 1963), Hilary Putnam's Twin Earth (Putnam 1975), Saul Kripke's 'Gödel' (Kripke 1980), John Searle's Chinese room (Searle 1980), and Derek Parfit's split brains (Parfit 1984)—to mention just a few well-known examples all became part and parcel of basic philosophical training, and appeals to them in arguments contributed to establish or to discredit various philosophical theses about, e.g., moral right and wrong, knowledge, meaning and reference, artificial intelligence, and personal identity. It would indeed be descriptively accurate to say that by the early Nineties thought experimentation had become deeply ingrained in analytic philosophy's self-conception, and judgments about hypothetical cases had grown to become the common stockin-trade of most mainstream debates in several areas of analytic philosophy (cf., e.g., Sorensen 1992; Häggqvist 1996). So much so, in fact, that many authors began to regard their use as "the distinctive method of contemporary analytic philosophy" (Horowitz and Massey 1991, p. 1).

Apart from a few notable, premonitory exceptions (cf., e.g., Fodor 1964; Wilkes 1988; Stich 1990), things began



Adriano Angelucci adriano.angelucci@uniurb.it

Margherita Arcangeli margherita.arcangeli@gmail.com

University of Urbino, Via Aurelio Saffi 2, 61029 Urbino, PU, Italy

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

to steadily, and rapidly change during the Nineties. In this regard, Michael DePaul and William Ramsey's 1998 influential collection of essays—Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry arguably constitutes a fundamental landmark in the recent history of analytic metaphilosophy, as many of the issues raised in that volume have contributed to shape, in one way or another, the currently thriving debate on philosophical method in general, and on the use of TEs in particular. In their Preface, the two editors stated—in a clearly programmatic spirit—that their main goal was that of initiating "a self-examination of philosophical method" which they believed was "long overdue" (DePaul and Ramsey 1998, Preface x). What made a careful, overall reassessment of standard philosophical methodology urgent, in their view, was the fact that analytic philosophers, for the most part, had failed to respond adequately to various challenges to the extant philosophical tradition that emerged from experimental research in cognitive psychology carried out over the previous two decades. As a matter of fact, DePaul and Ramsey were not alone in finding that the empirical evidence theretofore unearthed by studies of concepts and categorization judgments (cf., e.g., Smith and Medin 1981), on the one hand, and of inference strategies and belief revision (cf., e.g., Nisbett and Ross 1980), on the other, had a clear bearing on current philosophical practice. The results of these studies were indeed perceived by many philosophers as alarming, as they seemed to cast serious doubts both on the time-honored program of conceptual analysis as a reliable guide to philosophical truth, and on the widely relied upon method of reflective equilibrium as a rational way of forming beliefs. And to the extent that appeals to cases were at the heart of both, the use of TEs in philosophical theorizing was also called into question. A practice that was previously considered as an integral part of standard philosophical methodology, began to be regarded with suspicion in many circles. What appeared to be problematic about this practice, in particular, was its alleged reliance on intuitions as a trustworthy source of evidence about what would be the case ("what we would say") in a given hypothetical scenario, and hence as what ultimately drives philosophical inquiry. Indeed, as the title of DePaul and Ramsey's volume suggests, rethinking philosophical method, in the eyes of many, called for a systematic empirical investigation into the nature and functioning of these elusive mental states, as well as for a careful, and empirically informed assessment of their increasingly dubious epistemic credentials. "Western analytic philosophy"—as the two editors emphatically put it—"is, in many respects, undergoing a crisis where there is considerable urgency and anxiety regarding the status of intuitive analysis" (DePaul and Ramsey 1998, Preface, x).

Uneasiness about TEs began to spread, and around the year 2000 similar concerns about the status of "intuitive

analysis" ushered in one of the most thought-provoking and controversial developments in the recent philosophical landscape, i.e. the naturalistic-minded experimental turn brought about by the rise of experimental philosophy (henceforth X-Phi) (cf., e.g., Knobe and Nichols 2008, 2014; Horvath and Grundmann 2012; Alexander 2012; Sytsma and Livengood 2016). Broadly construed, X-Phi is a lively and multifaceted philosophical movement which—drawing on the methods of the social and cognitive sciences—regards as its general objective "the systematic collection and analysis of empirical data to help answer philosophical questions or solve philosophical problems" (Sytsma and Livengood 2016, p. 18). Work in X-Phi falls under what may be regarded as different yet mutually intersecting programs, with many practitioners in the field often participating in and many results bearing on—more than just one of them. Yet—despite the existence of a distinctively non-intuitional program in X-Phi (cf., e.g., Nichols 2002; Livengood et al. 2010)—most research carried out under this banner to date has either been concerned with understanding the nature of philosophical intuitions—i.e. addressing descriptive questions about what they are and how they work (cf., e.g. Greene et al. 2001; Knobe 2003)—or with assessing their seemingly problematic epistemic status—i.e. raising prescriptive questions about their evidential value (cf., e.g., Weinberg et al. 2001; Machery et al. 2004). The extant, and growing body of experimental results—typically, though not always, obtained by means of survey methods—appears to show that our allegedly intuitive judgments about many philosophical cases are worryingly sensitive to various kinds of truth-irrelevant factors, such as, e.g., demographic variables, situational cues, or the order in which the cases are presented (cf. Machery 2017, Ch. 2 for a detailed review of the relevant literature). According to many, these findings pose a serious *prima facie* challenge to the method of cases, and to the extent that this traditional practice plays a vital role within most current philosophical theorizing—i.e. that judgments about cases figure prominently in arguments aimed at supporting or undermining philosophical viewscarefully assessing its methodological soundness definitely looks like an intellectual endeavor well worth pursuing. Yet experimental methods and data, in and of themselves, can of course be put to different uses. In this regard, a common division within X-Phi work is the one between a negative and a positive program. Whereas the former is overtly pessimistic about the prospects of the method of cases, and recommends suspending judgment on the conclusions of though-experimentally based arguments (cf. Weinberg 2007; Alexander et al. 2010), the latter is more optimistic, and tends to conceive of X-Phi work as a valuable—perhaps indispensable—tool that would allow philosophers to improve on the traditional practice (cf., e.g., Knobe 2003; Buckwalter 2010). Regardless of which of the two programs



(if any) will prevail in the end, it seems particularly relevant, in the context of the present Introduction, to point out what is arguably a general tendency within the recent debate sparked off by the X-Phi challenge. It indeed seems fair to say that the primary focus of X-Phi-fuelled methodological concerns has somewhat shifted lately from the use of *intuitions* within the method of cases to the *method of cases* in and of itself, i.e. regardless of whether our allegedly intuitive judgments about cases actually rest on this otherwise mysterious intellectual faculty. Eduard Machery, one of the acknowledged pioneers of the field, has recently been explicit in this regard: "Experimental philosophers"—he writes—"have given the impression that their argument was directed at the alleged use of intuitions in philosophy. It is not; the target is the method of cases" (Machery 2017, p. 178).

If, as we believe, Machery is right in maintaining that the currently burning methodological questions raised by the abundant use of TEs in philosophical theorizing do not primarily concern—nor are in any way limited to—the purportedly fundamental role played by intuitions within this practice, then it seems that our prospects for making progress in the attempt at answering them will be dependent on a much broader understanding of that very practice. It is our hope that the selection of essays included in the present Issue will contribute to the achievement of this goal. We conclude this introduction with a brief summary of each contribution.

1 Synopsis

The idea of the present Issue originated in a workshop held at the University of Urbino, Italy, in June 2014, and subsequently developed into an independent editorial project by including contributions that were not initially presented at the workshop. The eight essays that follow—authored by young and emerging philosophers as well as fully accomplished ones-touch upon various aspects of the most recent debate surrounding TEs, closely engaging with many influential proposals that have been put forward over the last few years. They can be organised around three main areas of concern. The first group deals with general issues raised by the use of TEs, such as the possibility of drawing a principled distinction between scientific and philosophical TEs (Marco Buzzoni), the role played by intuitions within philosophical TEs (Julia Langkau), and the different ways in which the intended conclusion of a TE can be resisted (Roy Sorensen). The second group focuses on the logical form of philosophical TEs (Sören Häggqvist; Daniele Sgaravatti), and on the kind of access to modal knowledge they are standardly taken to provide (Vittorio Morato). The third and last group explores the virtues and limits of TEs used in two specific philosophical areas, namely the philosophy of science (Margherita Benzi), and the philosophy of mind (Elizabeth Schier).

1.1 The Nature of Philosophical Thought Experiments

In his "Thought Experiments in Philosophy: A Neo-Kantian and Experimentalist Point of View", Marco Buzzoni addresses the vexed issue concerning the relationship between philosophical and scientific TEs. Leaving aside skeptical concerns targeting the use of TEs per se, the current literature, according to Buzzoni, would be divided in two main camps: the optimists about the role of TEs in both science and philosophy, and the pessimists about the use of TEs in philosophy only. The pessimism of the latter camp, he argues, is fuelled by the idea that philosophical TEs—just as philosophical inquiry in general—would be less rigorous than their scientific counterparts due to their heavy reliance on misleading intuitions as well as on an utterly unconstrained use of imagination, as opposed to hard data and empirically informed hypotheses. Buzzoni disagrees with both camps, as, in his view, they would both be culpable of acknowledging only a difference in degree between the methods of science and those of philosophy—a stance he dubs "radical naturalism". He develops what he regards as an experimentalist and Neo-Kantian account of philosophical TEs. In a Kantian spirit, he argues in favour of a principled distinction between philosophical and scientific TEs. The experimentalist component, on the other hand, consists in maintaining that thought experimentation should not be confined to a form of purely armchair theorizing, as philosophy, in his view, needs science as much as the latter needs the former.

As we saw above, most critics of philosophical TEs pivot on their allegedly being "intuition pumps". Based on experimental findings, X-Phi practitioners have questioned the methodological soundness of this practice. Two general lines of reply have been pursued so far. The first one consists in defending the evidential role of intuitions, whereas the second one flat-out denies it. In her "Metaphilosophy and The Role of Intuitions", Julia Langkau focuses on two prominent instances of the second strategy, due to Max Deutsch, and to Herman Cappelen, and argues that both lines of reply to the X-Phi challenge fail. Within their texts, philosophers often provide arguments intended to support their claims about specific cases. According to Deutsch, this would be evidence that philosophers typically do not rely on intuitions in order to support their views. Langkau finds this unconvincing, and argues that reliance on arguments in philosophical texts is not incompatible with reliance on intuitions as evidence. Cappelen denies that philosophers typically appeal to something called 'intuitions' as evidence, and provides specific features that, in his view, we should look for in a



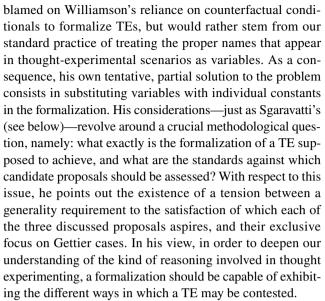
766 A. Angelucci, M. Arcangeli

philosophical text in order to decide whether such mental states are being appealed to as evidence in it. According to Langkau, this method would be unfit to deliver the goods, as philosophers' reliance on intuitions as evidence—as well as their metaphilosophical commitments more generally—may not be evident from their writings. In her view, the question of whether intuitions play a crucial role within the standard practice of appealing to TEs can only be answered within the broader framework of a general theory of knowledge, evidence, and philosophical methodology.

There are many ways of coping with and reacting to a philosophical TE. Roy Sorensen's contribution-"Smartfounding: Four Grades of Resistance to Thought Experiment"—introduces the new and intriguing notion of smartfounding, which he sees as the most sophisticated and virtuous way of resisting the intended conclusion of a TE, and contrasts with three other common, yet purportedly fallacious, reactions, namely the "unschooled response", dumbfounding, and imaginative resistance. Due to her lack of training, the unschooled would be simply bad at handling hypotheticals, especially when far-fetched, and hence perceived as having little grip on reality. She will typically either contradict what is stipulated or add information beyond the one explicitly given. Unlike her, the trained "dumbfounder" would be normally competent at thought experimenting, yet fallaciously selective in her willingness to follow the instructions given by the thought experimenter. In particular, she will commit performance errors when considering TEs which challenge her emotionally charged commitments, thereby lapsing back into the unschooled response. The third way of reacting to a TE, imaginative resistance, while not itself resting on performance errors, would be equally fallacious, as it will also typically eventuate in a failure to comply with the instructions received. As opposed to the other responses, the one offered by the "smartfounder"—and illustrated through a host of hilarious and thought-provoking examples—constitutes a very instructive and penetrating internal critique of TEs. The smartfounder practices "subversive compliance", and obeys instructions in a way that spoils the aim of the TE.

1.2 The Logic of Philosophical Thought Experiments

In his "Thought Experiments, Formalization, and Disagreement", Sören Häggqvist engages closely with three influential proposals—due to Timothy Williamson, Jonathan Ichikawa and Benjamin Jarvis, and Anna-Sara Malmgren, respectively—concerning the logical form of TEs understood as test benches for philosophical theories, and puts forward a proposal of his own, whose main formal feature is that of making use of free logic. In particular, he maintains that the much discussed problem of "deviant realizations"—contrary to what has been suggested—is not best



Daniele Sgaravatti is also interested in clarifying the peculiar kind of reasoning which characterizes any competent performance at thought experimenting. In his view, the main purpose of formalizing a TE would be that of illuminating the justificatory structure of the corresponding argument. In particular, the basic requirement that any acceptable account of the logical form of TEs should satisfy, according to him, would be that of containing a conditional. He further argues that we have some reasons to believe that the conditional in question should be a counterfactual. His contribution—"The 'If' in the 'What If""—is indeed a defence of Williamson's counterfactual account of TEs against two competing, and already mentioned proposals, due to Ichikawa and Jarvis, and Malmgren. According to Ichikawa and Jarvis, the problem of deviant realizations raised by Williamson's account can be taken care of by enriching the content of the thought-experimental scenario. Sgaravatti's reply consists in showing that the same strategy could also be adopted within Williamson's framework. On Malmgren's account, a complex possibility claim would exhaust our thought-experimental reasoning. Yet, according to Sgaravatti, leaving out the conditional would fail to capture an important part of our reasoning about cases, namely the way in which our judgment that we have a counterexample to the target thesis depends on our judgment that the situation described in the scenario is possible. In articulating his objections to both accounts, Sgaravatti draws attention to, and elaborates upon, what he sees as the cognitive and epistemic value of deviant scenarios.

Philosophical exceptionalists claim that our knowledge of metaphysical necessities and possibilities would be obtained by means of a special, a priori faculty, usually understood as conceivability. Anti-exceptionalists deny this. Williamson takes his counterfactual account to support anti-exceptionalism, as on this account the relevant modal knowledge would



be a by-product of our ordinary capacity to handle counterfactual conditionals (i.e. to develop counterfactual suppositions). In his "Conceivability, Counterfactual Thinking, and the Philosophical Exceptionality of Modal Knowledge", Vittorio Morato tries to show that Williamson's account is structurally similar to the more traditional conceivabilitybased account. In particular—he argues—according to both accounts the source of our relevant modal knowledge would be a general capacity to detect contradictions within suppositions. A consequence of this fact, in his view, is that Williamson's account is bound to face some of the same problems which seem to plague the conceivability-based one. The main drawback, according to Morato, would be that on this account every metaphysical necessity and possibility comes out as knowable (and unwelcome result he dubs "the problem of collapse"). The conclusion Morato draws is that Williamson's account, contrary to what is generally thought, is still based on a quite exceptional epistemic capacity to assess the truth of modal metaphysical claims which would correspond, on a metaphysical level, to a non-realist conception of metaphysical modality.

1.3 Philosophical Thought Experiments in Practice

Extant formal approaches to actual causation couched in terms of structural models tend to be regarded with suspicion by scientists and philosophers of science alike, due to the allegedly "unscientific" way in which philosophers typically theorize about it, namely by means of TEs. As consequence, one could be tempted to dismiss this traditional methodological approach, and to rely solely on its more rigorous counterpart in developing theories of the phenomenon under scrutiny. In her "Thought Experiments and Actual Causation", Margherita Benzi argues that this move would be unwarranted, as, in her view, TEs and structural models valuably complement each other. On the one hand she argues—structural models can profitably assist us both in deriving correct conclusions from a given definition, as well as in assessing whether a given imaginary case in fact constitutes a genuine counterexample to it. On the other, she contends that TEs, far from being "a relic of a bygone age", can usefully guide us in the building and assessment of structural models. The latter claim is brought out by showing how the so called problem of model isomorphism-generated by the possibility that two imaginary cases, while being represented by the same model, elicit strongly contrasting intuitions about what causes what—undermines the programmatic assumption according to which structural models would be sufficient to develop a satisfactory theory of actual causation. Contrary this assumption, she defends the view that structural models need to be integrated with some suitable, and context-sensitive, notion of normality, which TEs can help us identify.

In her "Subjectivity, Multiple Drafts, and the Inconceivability of Zombies and the Inverted Spectrum in This World", Elizabeth Schier takes issue with the proponents of the hard problem of consciousness, who take zombie and inverted spectrum TEs—i.e. situations in which we imagine all physical truths about the world holding and yet the phenomenal truths vary—to show that consciousness cannot be physical. The gist of her argument is that, given that materialism is a claim about the nature of our world, the difficult epistemic task faced by David Chalmers' conceivability argument for dualism would be that of making sure that we are not ignorant of the physical and psychological facts that make zombies and inverts inconceivable in this world. She further argues that, by understanding the limitations put on our cognitive systems by the physics of our world, one can come to see that zombies and inverted spectra are not ideally conceivable. She tries to establish this point by arguing that the "Cartesian Theatre" model of consciousness implicit in both zombie and inverted spectrum TEs is false of humans, and that, given the time constraints our cognitive systems need to work under, there is good reason to think that it must be false of all cognitive systems in our universe. On her way to this conclusion, she draws a parallel with Putnam's Twin Earth. The reason why this TE does not undermine materialism for water, in her view, would be that, given the physics of our world, and provided one understands the chemistry of water, it is ideally inconceivable that water could be made of anything other than H₂O.

References

Alexander J (2012) Experimental philosophy. An introduction. Polity, Cambridge

Alexander J, Mallon R, Weinberg J (2010) Accentuate the negative. Rev Philos Psychol 1(2):297–314

Buckwalter W (2010) Knowledge isn't closed on Saturday: a study in ordinary language. Rev Philos Psychol 1:395–406

Cappelen H (2012) Philosophy without intuitions. Oxford University Press, Oxford

Cappelen H, Szabó Gendler T, Hawthorne J (2016) The Oxford handbook of philosophical methodology. Oxford University Press,

D'oro G, Overgaard S (eds) (2017) The Cambridge companion to philosophical methodology. Cambridge University Press, Cambridge DePaul M, Ramsey W (eds) (1998) Rethinking intuition: the psychology of intuition & its role in philosophical inquiry. Rowan and Littlefield, New York

Deutsch M (2015) The myth of the intuitive. MIT Press, Cambridge Fischer E, Collins J (2015) Experimental philosophy, rationalism, and naturalism. Rethinking philosophical method. Routledge, London Fodor J (1964) On knowing what we would say. Philos Rev 73:198–212 Foot P (1967) The problem of abortion and the doctrine of the double

Gettier E (1963) Is justified true belief knowledge? Analysis 23:121-23

effect in virtues and vices. Oxford Rev 5:5-15



768 A. Angelucci, M. Arcangeli

Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. Science 293/5537:2105–2108

- Häggqvist S (1996) Thought experiments in philosophy. Almqvist & Wiksel. Stockholm
- Haug M (2014) Philosophical methodology: the armchair or the laboratory? Routledge, London
- Horowitz T, Massey G (eds) (1991) Thought experiments in science and philosophy. Rowman and Littlefield, Savage
- Horvath J, Grundmann T (2012) Experimental philosophy and its critics. Routledge, London
- Jackson F (1998) From metaphysics to ethics. A defence of conceptual analysis. Oxford University Press, Oxford
- Knobe J (2003) Intentional action and side-effects in ordinary language. Analysis 63:190–194
- Knobe J, Nichols S (eds) (2008) Experimental philosophy. Volume I. Oxford University Press, Oxford
- Knobe J, Nichols S (eds) (2014) Experimental philosophy. Volume II. Oxford University Press, Oxford
- Kripke S (1980) Naming and necessity. Blackwell, Oxford
- Livengood J, Sytsma J, Feltz A, Scheines R, Machery E (2010) Philosophical temperament. Philos Psychol 33:313–330
- Machery E (2017) Philosophy within its proper bounds. Oxford University Press, Oxford
- Machery E, Mallon R, Nichols S, Stich S (2004) Semantics, cross-cultural style. Cognition 92:B1–B12
- Nado J (2016) Advances in experimental philosophy and philosophical methodology. Bloomsbury Academic, New York

- Nichols S (2002) On the genealogy of norms: a case for the role of emotion in cultural evolution. Philos Sci 69:234–255
- Nisbett R, Ross L (1980) Human inference: strategies and shortcomings of social judgment. Prentice-Hall, Englewood Cliffs
- Parfit D (1984) Reasons and persons. Clarendon Press, Oxford
- Putnam H (1975) Meaning of "Meaning". Minnesota Stud Philos Sci 7:131–193
- Searle J (1980) Minds, brains, and programs. Behav Brain Sci 3:417–445
- Smith E, Medin D (1981) Categories and concepts. Harvard University Press, Cambridge
- Sorensen R (1992) Thought experiments. Oxford University Press, Oxford
- Stich S (1990) The fragmentation of reason. MIT Press, Cambridge Stich S, Tobia K (2016) Experimental philosophy and the philosophical tradition. In: Sytsma J, Buckwalter W (eds) A companion to experimental philosophy. Wiley, New York, pp 5–21
- Sytsma J, Livengood J (eds) (2016) The theory and practice of experimental philosophy. Broadview Press, Peterborough
- Weinberg J (2007) How to challenge intuitions empirically without risking skepticism. Midwest Stud Philos 31:318–343
- Weinberg J, Nichols S, Stich S (2001) Normativity and epistemic intuitions. Philos Topic 29:429–460
- Wilkes KV (1988) Real people. Personal identity without thought experiments. Clarendon Press, Oxford
- Williamson T (2007) The philosophy of philosophy. Blackwell, Malden

