

Projects and Methods of Experimental Philosophy

Eugen Fischer and Justin Sytsma

How does experimental philosophy address philosophical questions and problems? That is: What projects does experimental philosophy pursue? What is their philosophical relevance? And what empirical methods do they employ? Answers to these questions will reveal how experimental philosophy can contribute to the longstanding ambition of placing philosophy on the ‘secure path of a science’, as Kant put it. We argue that experimental philosophy has introduced a new methodological perspective – a ‘meta-philosophical naturalism’ that addresses philosophical questions about a phenomenon by empirically investigating how people think about this phenomenon. This chapter asks how this novel perspective can be successfully implemented: How can the empirical investigation of how people think about something address genuinely philosophical problems? And what methods – and, specifically, what methods beyond the questionnaire – can this investigation employ? We first review core projects of experimental philosophy and raise the question of their philosophical relevance. For ambitious answers, we turn to experimental philosophy’s most direct historical precursor, mid-20th century ordinary language philosophy, and discuss empirical implementations of two of its research programmes that use experimental methods from psycholinguistics and corpus methods from the digital humanities.

1. Introduction

Philosophers have long wondered how their discipline can be placed on the ‘secure path of a science’, with methods that go beyond ‘groping among mere concepts’ (cf. Kant 1787/2003, p.21). Much of the excitement attaching to experimental philosophy arises from the fact that it holds out the promise of fresh answers to this question. Experimental philosophy recruits empirical methods to address philosophical questions and problems. The empirical methods are taken mainly from the social and the cognitive sciences, although increasingly also from linguistics and the digital humanities. Many of these methods are impeccably scientific and reasonably well understood, and certainly go beyond ‘mere groping’. To answer the Kantian question, though, we need to get clear on how these methods can be deployed for *specifically* philosophical projects, and this question is itself philosophically difficult and divisive.

Experimental philosophy has already addressed questions from almost all areas of philosophy, with a focus on ethics (26% of current experimental philosophy papers, as categorised by the PhilPapers database), the philosophy of action (12%), epistemology (10%), philosophy of language (8%), philosophy of mind (7%), and metaphysics (6%).¹ Within each area, experimental literatures have developed around specific topics, such as free will and intentional action in the philosophy of action, causation and personal identity in metaphysics, or contextualism in epistemology. Some of these foci have emerged from philosophically unexpected findings in seminal studies – such as, for example, the finding that judgments traditionally regarded as merely descriptive or explanatory, such as attributions of intentionality or causal efficacy, are influenced by the moral valence of the relevant outcomes (Knobe 2003, Nadelhoffer 2006, Knobe and Fraser 2008). Other foci have emerged from well-

¹ These figures are based on PhilPapers listings, as of November 9th, 2021. They indicate lower bounds, since listings under the generic ‘Experimental Philosophy’ heading include more non-experimental papers (e.g., on foundations of experimental philosophy) than sub-categories. (e.g., ‘Experimental Philosophy of Action’).

established theoretical debates that explicitly rely on straightforwardly empirical assertions. For instance, epistemological debates about contextualism rely on claims about how contextual factors influence ordinary speakers' attributions of knowledge (DeRose 1992, Buckwalter 2010, Schaffer and Knobe 2012). Epistemologists advanced competing claims to account for the same intuitions about the same hypothetical cases. This clearly called for empirical investigation of competing explanations. However, the dynamic growth of experimental philosophy – from under 100 papers in 2007, to over 700 papers listed under this heading in PhilPapers in April 2014, to more than 2,750 papers listed in March 2022 – has seen topical coverage extend well beyond these foci of interest (see subsequent chapters in this volume; also Sytsma and Buckwalter 2016). Given this diversity in topics, projects in experimental philosophy in general can be usefully discussed only at the meta-philosophical level.

At the grandest meta-philosophical level, experimental philosophers are united by a commitment to methodological naturalism. Naturalism as a methodological stance is independent from naturalism as a metaphysical position (De Caro 2008, Haug 2014, pt.1). Methodological naturalism recommends we address philosophical problems by drawing on empirical findings from the sciences (whose findings need not be consistent with metaphysical naturalism; Collins 2015). Traditional methodological naturalism seeks to address philosophical problems about a topic (say, vision, or intentional action) by drawing on empirical findings about that topic (say, findings from the psychology or neuroscience of vision, or of action control). Experimental philosophers go further by conducting their own empirical investigations when called for, distinguishing their practice from empirically informed 'empirical philosophy' (Prinz 2008).

Crucially, however, experimental philosophers are conducting empirical investigations with two very different directions of thrust, one of which is genuinely novel. To address, for example, philosophical questions or problems about vision, experimental philosophers have not only built on the psychology and neuroscience of vision and added their own investigations of the mechanisms of vision (e.g., Schwenkler and Weksler 2019, Weksler, Jacobson, and Bronfman 2021). Rather, experimental philosophers of perception have also built on the psychologies of *judgment and language*, and have empirically investigated how people *think and speak about* vision (e.g., Fischer et al. 2021, Fischer, Engelhardt, and Sytsma 2021, Roberts, Allen, and Schmidtke 2018, 2021). Similarly, to address philosophical questions about colour, experimental philosophers have primarily investigated not colour perception but colour *cognition* (e.g., Cohen and Nichols 2010, Sytsma 2010, Hansen and Chemla 2017, Roberts and Schmidtke 2019), and similarly with regard to other topics, such as pain (e.g., Sytsma and Reuter 2017, Liu 2020, Reuter and Sytsma 2020, Salomons et al. 2021). Where traditional ('object-level') methodological naturalism addresses philosophical questions and problems about a topic (e.g., vision or colour) by drawing on scientific findings about that topic, experimental philosophy has added a novel ('meta-level' or) *meta-philosophical naturalism* that addresses those questions and problems through empirical study of *how people think about* the topic of interest (Fischer and Collins 2015).

This distinctive new stance motivates our guiding question: How can empirical findings about how people think about a topic be used to address philosophical questions and problems about that topic? Our chapter will develop this question and make a start on answering it. We will develop and address the question through a review of experimental philosophy's two major strands – evidential and explanatory experimental philosophy (Section 2). We then propose fresh answers through two case studies (Sections 3–4).

The case studies explore how recent work in experimental philosophy has developed and transformed research programmes initially formulated by experimental philosophy's most direct historical precursor: Ordinary language philosophy, as pioneered by J.L. Austin (1957, 1962) and – on a liberal understanding of the movement – Arne Naess (1956, 1961), was analytic philosophy's first attempt to overcome limitations of armchair reflection through the use of (informal) experiments (Hansen and Chemla 2015), (peer-based) focus groups (Urmson 1969), and empirical surveys (Murphy 2014). This precursor of experimental philosophy² suggests answers to our guiding question: Ordinary language philosophy adopted a similar 'meta-level' perspective and deployed insights into *how people talk about* a topic in order to address philosophical questions and problems about that topic. The two main strands of ordinary language philosophy explored different approaches in this vein. To answer our guiding question, our case studies will therefore consider how experimental philosophy has developed, respectively, the programs of *critical* ordinary language philosophy (Section 3) and *constructive* ordinary language philosophy (Section 4). So far, experimental philosophy has relied mainly on questionnaire-base methods (explained in Sytsma and Livengood 2015), while further methods are increasingly being taken up from two sources: behavioural experiments from the social sciences and empirical and computational methods from the digital humanities (see Fischer and Curtis 2019). Our case studies will illustrate new uses, in experimental philosophy, of such further methods, namely, experimental methods from psycholinguistics (Section 3) and corpus methods from the digital humanities (Section 4).

2. Experimental Philosophy: Evidential and Explanatory

While experimental philosophy ranges considerably more widely, its two most prominent strands examine philosophically relevant intuitions (Fischer and Collins 2015, Sytsma and Livengood 2015). In several areas of philosophy, theory construction involves the 'method of cases': In thought experiments, philosophers consider (typically) hypothetical cases (like Gettier cases, trolley cases, Frankfurt cases, and so forth), elicit intuitive judgments about them, and deploy these judgments as defeasible evidence for or against philosophical theories. These theories are typically required to be consistent with the case judgments and are often meant to explain or justify them. This approach has been used to assess analyses of concepts (e.g., 'knowledge'), modal claims (e.g., that, provided current chemistry is correct, water must be H₂O), and moral claims (e.g., the moral equivalence of killing people or letting them die), among others (for reviews, see Daly 2010, Machery 2017, pp. 11–44). Philosophical theory construction guided by the method of cases often proceeds by working back and forth between intuitions elicited by different thought experiments and various background beliefs, until 'reflective equilibrium' and a coherent set of judgments and beliefs has been achieved. While used most explicitly in moral and political philosophy (e.g., Foot 1967, Rawls 1971), this approach is sufficiently common to have been characterised as analytic philosophy's 'standard justificatory procedure' (Bealer 1996, p. 4). Indeed, some philosophers have regarded it as 'characteristic, perhaps definitive, of philosophical argumentation throughout its history' (Levin 2005, p. 193). Methodological rationalists have invoked the approach, as practiced from the armchair, to argue for the disciplinary autonomy of philosophy as a subject, holding that it can gain insight through a priori methods that rely on intuition and pure reason alone (e.g., Bealer 1996, 2000).

² See Chapter 1 for discussion of further precursors of experimental philosophy.

The main strands of experimental philosophy are responses to such intuition-driven philosophising: *Evidential experimental philosophy* seeks to assess the evidential value of philosophically relevant intuitive judgments ('philosophical intuitions'), while *explanatory experimental philosophy* seeks to explain such intuitions.³ The former may appear to make at most indirect contributions to philosophical theorising; the latter might seem to be of primarily psychological interest. In the following, we examine the philosophical ambitions and rationales of these two enterprises.

2.1 Evidential Experimental Philosophy

Evidential experimental philosophy is pursued with two different ambitions (for reviews, see Stich and Tobia 2016, Mallon 2016) – what are often labelled the *positive* and *negative* programmes. We begin with the *positive programme*, which shares the key ambitions and methodological commitments of the pre-experimental philosophical work it engages with and introduces empirical methods to better sort the evidential wheat from the chaff.

Many philosophical debates take for granted the existence of 'common-sense' conceptions of the phenomena of interest, and accord these conceptions an epistemic default status that places the burden of proof on critics (e.g., Daly 2010, Jackson 1998, Strawson 1959). This assumption is common, for example, in debates about free will (e.g., O'Connor and Franklin 2021, Jackson 1998, p.31), consciousness (e.g., Chalmers 1996, Lewis 1972), colour (e.g., Allen 2016, Johnston 1992), and time (e.g., Callender 2017, Price 2011), among others. Seminal contributions to the positive programme sought to verify this assumption, as made in debates about free will. In these debates, common-sense status has been claimed for both compatibilist and incompatibilist conceptions of free will (see Chapter 11). Nahmias and colleagues (2004, 2005, 2006) examined whether laypeople's intuitive attributions of moral responsibility to hypothetical agents in deterministic and non-deterministic universes, respectively, are consistent with these conceptions. Other studies have examined lay conceptions of consciousness (e.g., Sytsma and Machery 2010; see Chapter 10), colour (e.g., Roberts and Schmidtke 2019), or time (e.g., Latham, Miller, and Norton 2021), among others. A prominent motivation of these studies has been to settle which of competing conceptions is the clear majority view among adult laypeople and can lay claim to being 'common sense' and enjoying epistemic default status. Viewed thus, these studies adjudicate on burdens of proof.

Other contributions to the positive programme have sought to enhance the use of the method of cases in conceptual analysis – e.g., on the concepts of *knowledge* (Buckwalter 2010, Schaffer and Knobe 2012; see Chapter 5) or *intention* (for a review see Cova 2016). Much work of this kind accepts the method's fundamental assumption that 'possessing a concept makes one disposed to have pro-intuitions toward correct applications and con-intuitions toward incorrect applications – correct, that is, relative to the contents of the concept as it exists in the subject's head' (Goldman 2007, p. 15, cf. Chalmers and Jackson 2001, Ludwig 2007) and seek to arrive at characterisations of conceptual contents through abductive inferences from (patterns of) intuitive application judgments. But this body of work questions the related assumption that individual philosophers, simply in virtue of being competent speakers of a shared language like English, can effectively examine folk concepts by eliciting just their own

³ See Sytsma and Livengood (2015). We employ the more expressive label 'explanatory experimental philosophy' for what they termed the 'cognitive program'. On this first rough characterization, the programs overlap: Many (e.g., debunking) explanations are developed to contribute to epistemological evaluation.

intuitions.⁴

Some critics of standard uses of the method of cases have suggested that philosophers' judgments may be unduly influenced by philosophical *theories* of the concepts of interest (*knowledge, moral permissibility, intention*, and so forth), rather than accurately reflecting folk concepts (e.g., Goldman 2007, Weinberg, Nichols, and Stich 2001). This suggestion motivates turning to case judgments from laypeople. This turn is warranted by findings that the intuitive judgments of philosophers sometimes diverge systematically from the judgments of laypeople or even other academics (e.g., Starmans and Friedman 2020 on intuitive knowledge attributions; cf. Horvath and Wiegmann 2016), and by patterns of responses from laypeople that philosophers failed to anticipate (e.g., Knobe's (2003) discovery that intention attributions are influenced by the moral valence of actions' outcomes).⁵ Thus, the traditional assumption that philosophers' case intuitions are representative of competent speakers' intuitions more generally and can be safely taken to reflect widely shared folk concepts clearly stands in need of empirical investigation for any given case of interest. Empirical studies examining this assumption help determine whether philosophers' case judgments provide relevant evidence for the analysis of folk concepts and provide relevant evidence where philosophers' own judgments won't do.

By contrast, the *negative programme* in evidential experimental philosophy puts the method of cases more directly to the test by examining how epistemically irrelevant factors influence intuitions. To spell out the challenge raised by the negative programme, it is worth distinguishing between two different uses of the method of cases: in the 'formal mode' the method is used to analyse concepts; in the 'material mode' it is used to establish truths about the world. In different ways, both uses are put into doubt by the sensitivity of case judgments to irrelevant factors.

The 'formal' use assumes that case judgments reflect the content of the underlying concepts. This assumption will not hold if case judgments are notably influenced by factors that have nothing to do with the content of the concept at issue. Factors examined include presentational factors such as purely verbal differences between equivalent case descriptions, the order in which cases are presented, or physical circumstances under which cases are presented (e.g., Horvath and Wiegmann 2021; for reviews, see Machery 2017, Stich and Tobia 2016). They further include individual differences in personality traits or cognitive traits like reflectiveness that may influence how a speaker applies a concept to a case, irrespective of the concept's represented content. Differences between demographic groups can raise similar worries. Sensitivity to cultural or socio-economic background can raise doubts about whether the use of a concept of interest is universal and have us ask whether different communities employ different concepts and, if so, which of these concepts is better, for which purposes (Stich and Tobia 2016, p. 14). Such findings move philosophical research on the concepts of interest from the traditional concerns of conceptual analysis to the ameliorative concerns of conceptual engineering (for a review see Cappelen and Plunkett 2020). Experimental philosophy can then contribute to this new enterprise in various ways – for example, by examining to what extent people are prone to reason badly with extant concepts (which are

⁴ A related body of research on folk concepts is more profoundly sceptical of the method of cases and has employed psycholinguistic methods, instead (perhaps first: Powell, Horne, and Pinillos 2014, for a recent example, see Beisbart and Reuter 2021).

⁵ Machery and Stich (2022, Table 2) list twenty studies contributing to conceptual analysis that elicit lay judgments that arguably differ from the case intuitions capturing or close to a 'textbook consensus'.

therefore in need of improvement) or are able to reason correctly, or as intended, with supposedly improved concepts (e.g., Fischer 2020, Machery 2021).

The ‘material’ use of the method of cases assumes that intuition is a reliable source of the case judgments of interest. The basic assumption is that these judgments are (at least) true more often than not. This is meant to warrant the further assumption that they are generated by an underlying cognitive process that is reliable. If so, the fact that we have these intuitions, as and when we do, speaks for their truth. The basic assumption is called into doubt by sensitivity to presentational factors and individual differences as well as by sensitivity to demographic factors: If any of these factors lead to different judgments, the judgments made under some conditions, or by some people, must be wrong. In the absence of error theories that allow us to understand when and why people make the wrong judgments, or under what conditions and in what kind of people the underlying process is unreliable, however, we cannot tell which intuitions provide evidence for claims about the world.

The negative programme thus examines basic assumptions that motivate different uses of the method of cases, for individual classes of intuitive case judgments. Findings help philosophers identify classes of intuitions that cannot serve as evidence for or against philosophical hypotheses – or at any rate, cannot serve as such evidence in the absence of error theories and, more generally, a better understanding of the underlying processes. This allows philosophers to restrict their use of the method of cases to kinds of judgments that are not otiose due to inexplicable sensitivity.

To sum up, evidential experimental philosophy stands to make a variety of contributions to philosophy: its findings help to adjudicate on burdens of proof, to prevent premature generalization from philosophers’ intuitions, and to avoid reliance on dubious evidence and methods incapable of supporting or assessing claims of interest. While the positive programme also generates fresh evidence about common-sense conceptions and folk concepts, both the positive and the negative programme in evidential experimental philosophy contribute mostly at the meta-level, where they address philosophical questions about the phenomenon of interest indirectly, by answering methodological questions about how those philosophical questions should be addressed, and what evidence should be admitted.

2.2 Explanatory Experimental Philosophy

Many contributions to evidential experimental philosophy involved explanations of the intuitions at issue – e.g., on free will (Nichols and Knobe 2007, Nadelhoffer et al. 2020; see Chapter 11), consciousness attributions (Arico et al. 2011, Fischer and Sytsma 2021; see Chapter 10), or knowledge attributions (Alexander, Gonnerman, and Waterman 2014, Gerken et al. 2020; see Chapter 5). In this work, explanation (debunking and other) is at the service of epistemological evaluation.

Other work in experimental philosophy pursues explanatory aims for their own sake. This work can be regarded as giving rise to a distinct strand of experimental philosophy that is continuous with cognitive science. It follows the same pattern as cognitive science, where researchers seek to identify new effects and explain them in terms of cognitive processes, typically with an aim of explaining a maximum number of effects in terms of a minimum number of processes. This pattern is well-illustrated by work sparked by the side-effect effect. For example, Chandra Sripada (2010, 2012) proposed the ‘deep self’ hypothesis that explains this effect by positing a process whereby people attribute to agents a set of deeply held values and determine whether the agent’s action concords with this ‘deep self’. The hypothesis was

invoked from the start to explain not only the targeted effect on intention attributions, but also effects on assessments of moral responsibility (Sripada 2010) and freedom (Sripada 2012). The proposal then swiftly motivated further research on the process that examined its impact on yet further phenomena (starting with Newman, Bloom, and Knobe 2014).

Discerning the characteristic course of cognitive science research in much experimental philosophy, Joshua Knobe (2016, p. 50) concludes that ‘the vast majority of [experimental philosophy] research is *cognitive science*.’ We propose a more nuanced approach that takes into account how such explanations are intended to fit into the philosophical dialectic. Where explanatory efforts are motivated by the aim to contribute to the epistemological assessment of philosophical intuitions, claims, or arguments, we see such work as an extension of evidential experimental philosophy. By contrast, where the explanans (e.g., people’s ‘deep self’ conception) is treated as a phenomenon of philosophical interest in its own right, regardless of how explanations invoking it might bear on further (meta-)philosophical questions, we hold that it forms a distinct strand of experimental philosophy: *explanatory experimental philosophy*.

As it unfolds, a line of research may well, by these lights, change from evidential experimental philosophy to explanatory, and back again. Explanatory work initially motivated by an evidential question may come to be pursued for its own sake, and the insights gained in this pursuit might later be applied to a new set of evidential questions. For instance, as Knobe (2016) points out, work sparked by the side-effect effect that follows the trajectory of cognitive science works against the ambitions of conceptual analysis: As standardly conceived, conceptual analysis seeks to explain patterns of case judgments, first and foremost, by reference to the content of the concept of interest. Work following the cognitive science trajectory aims to explain those patterns by reference to cognitive processes that affect the use of several concepts. But the more substantive the theories of these underlying cognitive processes become, the less explanatory work remains to be done by accounts of individual contents. Work on theories of such cognitive processes is hence set to leave the ambit of conceptual analysis. At the same time, some such theories can in principle be repurposed for the negative programme: Where they expose biases – like the ‘blame bias account’ of the side-effect effect (Nadelhoffer 2006, Alicke 2008) – they can potentially provide a basis for error theories that allow us to understand when and why people make wrong judgments. Without such theories, the negative programme can infer from intuitional sensitivity only that certain kinds of intuitions are unreliable. With suitable error theories, it can go beyond purely negative findings and hope to identify conditions under which people may trust their intuitive case judgments (Weinberg 2015).⁶

Explanatory experimental philosophy thus seems to face a tougher question about its philosophical relevance than evidential experimental philosophy. The latter seemed to contribute to philosophical questions and problems mainly indirectly, by answering methodological questions about how those questions should be addressed, and what evidence should be admitted. Where explanatory efforts are put to these uses, they share this indirect relevance. By contrast, explanatory experimental philosophy, which undertakes explanatory work for its own sake, might be accused of doing something else entirely. Insofar as cognitive science is taken to be distinct from philosophy, accepting with Knobe that evidential

⁶ E.g., Fischer and Sytsma (2021) expose a framing effect in judgments about philosophical zombies and explain it by reference to a comprehension bias. This explanation allows to go beyond the conclusion that intuitions about philosophical zombies are unreliable, and identify a more and a less helpful frame.

experimental philosophy is cognitive science raises a critical question – ‘Why is this work *philosophy*?’ (Stich and Tobia 2016, fn. 10). At this point, intuitional experimental philosophy appears to face a dilemma: the critic can charge that it either contributes to philosophy only indirectly (through the evidential strand) or not at all (if the explanatory strand reduces to ‘mere psychology’).

3. Critical Ordinary Language Philosophy: Psycholinguistic Experiments

Much experimental philosophy seeks to address philosophical problems about a topic by examining how people think and speak about that topic. The apparent dilemma we now face motivates the question, how an approach along these lines can make a direct and substantive contribution to resolving distinctively or characteristically philosophical problems. For a first answer, we turn to experimental philosophy’s most direct historical precursor, ordinary language philosophy. Much work in this mid-20th century movement focused on ‘philosophical puzzles’ which, ordinary language philosophers maintained, arise from the way we think and speak about the phenomena that puzzle us (e.g., Austin 1946, 1962, Ryle 1954, Waismann 1997; cf. Fischer 2019, Schroeder 2006, pp. 151-168, Wittgenstein 1953). Understanding these ways – and their errors – was regarded as an essential step in resolving the resulting problems. We now consider such ‘puzzles’ (Section 3.1), spell out how experimental philosophy can provide much-needed empirical foundations for an approach to them that is in line with critical ordinary language philosophy (Section 3.2), and explain how experimental methods from psycholinguistics can be employed for the purpose (Section 3.3).

3.1 Philosophical Puzzles

Plato regarded a sense of wonder in the face of the familiar as the starting point of philosophising (n.d./2004, 155b–d). He had in mind puzzlement about *the very possibility* of familiar facts, rather than curiosity about their causes – not ‘what makes the apple drop?’ but ‘how is it even possible for the apple to reach the ground?’ (or the tortoise to overtake Achilles). This sort of puzzlement is the hallmark of *one* kind of characteristically philosophical problems, exemplified by sceptical problems and the problems of mental causation, free will, perception, and induction.⁷ These philosophical puzzles (let’s call them ‘Platonic puzzles’) are motivated by intuitive lines of thought that suggest certain common-sense convictions *cannot* be true. These lines of thought, often conceptualised as paradoxes, appear to bring out antinomies and motivate philosophical questions like: How is it possible to know anything at all about the physical environment? How can our beliefs and desires make a difference to our bodily movements? How is it possible that people are morally responsible for their actions? How is sense-perception, as we ordinarily understand it, even possible?

Relevant lines of thought typically rely on apparently common-sense convictions and intuitive judgments, and suggest a conclusion at odds with further common-sense convictions. These lines of thought thus appear to bring out conflicts between extant beliefs, namely, between the beliefs invoked by the arguments and the belief(s) challenged by them. To the extent to which these beliefs are deeply held, these conflicts are both puzzling and troubling, causing cognitive dissonance (see Gawronski and Strack 2012, Kruglanski et al. 2018).

⁷ The entries in the *Stanford Encyclopedia of Philosophy* on mental causation (Robb and Heil 2021), free will (O’Connor and Franklin 2021), the problem of perception (Crane and French 2021) and the problem of induction (Henderson 2020) all conceptualize at least key versions or components of these problems in a manner consistent with the following characterization.

Philosophical responses typically seek to show that these conflicts are merely apparent. Responses may include revisionary elements (some common-sense convictions get revised) and diagnostic elements (some steps of the arguments are shown fallacious). But, typically, they focus on reconciliation: on showing that, and explaining why, the common-sense conviction or conception challenged (or as much of it as survives revision) is consistent with as much of the arguments challenging them as theorists feel obliged to accept (see Fischer 2011 for a review).

Platonic puzzles are developed by, and perhaps even arise from, intuitive reasoning about the world. These puzzles therefore seem a particularly plausible target for the meta-philosophical naturalism promoted by experimental philosophy (Section 1): They should help us identify some ways in which empirical findings about how people think about a topic can be used to address philosophical problems about that topic. For a start, experimental philosophy can support and complement the responses reviewed:

- Experimental philosophy can help establish whether supposed common-sense convictions – invoked or apparently challenged – are indeed common sense. Similarly, it can assess whether people indeed make any such intuitive judgments as the arguments rely on.
- Going a step further, it can help *assess* these convictions and intuitions, e.g., by developing debunking explanations of the relevant beliefs and judgments. Together, these first two steps assess the need for revision of common sense and identify the parties in need of reconciliation.
- Finally, experimental philosophy can help assess the arguments that rely on those beliefs or judgments, for example, by exposing biases that lead to typically overlooked fallacies. The identification of fallacies supports diagnostic responses.

We now consider a case-study that develops the last and most novel of these applications and demonstrates how experimental philosophy can directly address Platonic puzzles.

3.2 A New Experimental Project: Critical Ordinary Language Philosophy

This case-study is on an experimental implementation of critical ordinary language philosophy, as pioneered by J.L. Austin (1946, 1962). This purely diagnostic approach seeks to ‘dissolve’ philosophical (‘Platonic’) puzzles by exposing fallacies already in the opening moves of the arguments that raise them; the aim is to reveal that these arguments fail to get off the ground, so that the apparent conflict between extant beliefs does not even begin to arise and is *merely* apparent. Following Austin (1962), the approach has been developed most fully for the ‘problem of perception’ (Crane and French 2021, Smith 2002). As standardly conceived today, this problem is developed by arguments ‘from illusion’ and ‘from hallucination’. These arguments invoke the uncontroversial beliefs that illusions occur and hallucinations are possible, and derive the conclusion that perceivers are directly aware of subjective sense-data, rather than physical objects. The arguments thus appear to challenge the common-sense conception of sense-perception (folk direct realism), and bring out a conflict between this conception and those uncontroversial premises. They motivate the question, how sense-perception, as we ordinarily understand it, is even possible – given the evident possibility of illusions and hallucinations.

Critical ordinary language philosophy seeks to expose ‘seductive (mainly verbal) fallacies’ as ‘concealed motives’ for formulating philosophical paradoxes (Austin 1962, p. 5). That the fallacious inferences are ‘concealed’ means that thinkers are not conscious of making

them and presupposing their conclusions, in the relevant arguments. In today's terms, the approach focuses on exposing automatic default inferences that are contextually inappropriate (Fischer et al. 2021). This approach is in need of experimental implementation and empirical validation: Thinkers have no privileged access to automatic inferences, which can be documented only experimentally (Bargh et al. 2012). Moreover, principles of charity demand that fallacies be attributed to competent thinkers only in the light of an empirically supported explanation of when and why competent thinkers commit such fallacies (Thagard and Nisbett 1983). An extension of evidential experimental philosophy that turns from the assessment of intuitions to that of inferences in verbal reasoning can provide – and has provided – just what is needed.

The need for a validating explanation is pressing for attributions of inappropriate default inferences: Nouns and verbs automatically activate associated stereotypes (prototypes and situation schemas, respectively) that represent knowledge about typical and diagnostic features of objects and events (Engelhardt and Ferreira 2016). These stereotypes support defeasible inferences (the 'secretary' is female) that are triggered by words almost irrespective of context (Levinson 2000). These inferences are swiftly complemented, as combinations of nouns and verbs ('the mechanic checked...') activate knowledge about the typical features of more specific situations (car inspections) that are not activated by either word alone (Bicknell et al. 2010, Matsuki et al. 2011). In sentence comprehension, these different inferences need to be integrated with each other and with information from wider discourse context (Metusalem et al. 2012), in building the *situation model*: the mental representation of the situation described by the text or speech, which provides the basis for further judgements and reasoning about that situation (Kintsch 1988, Zwaan 2016). Initial inferences decay when lacking contextual support (Oden and Spira 1983) and may get effortfully suppressed where they conflict with contextual information ('the secretary, Mr. Smith') or with background beliefs (Faust and Gernsbacher 1996). Competent language users are good at this task (for a review see Butterfuss and Kendeou 2018), and can, e.g., suppress default inferences that conflict with background beliefs within one second (Fischer and Engelhardt 2017). In consequence, inappropriate default inferences hardly influence further judgment and reasoning.

One explanation of when and why contextually inappropriate default inferences influence further cognition, even so, is provided by the linguistic salience bias hypothesis (Fischer and Engelhardt 2020, Fischer and Sytsma 2021): Subordinate uses of markedly unbalanced polysemes (like 'see') can give rise to fallacies of equivocation. That is: Subordinate uses ('Jack saw her point') are often interpreted by retaining the initially activated stereotype (situation schema) associated with the dominant sense (e.g., the visual sense of 'S sees X') and suppressing its irrelevant component features (e.g., *S looks at X, X is in front of S*, and so forth, retaining only *S knows X is there, S knows what X is*) (Giora 2003; cf. Brocher et al. 2018). Where some, but not all, of the stereotypical features associated with the dominant sense are contextually relevant, and the dominant sense is far more frequent than all others, general principles of activation ensure that suppression remains partial.⁸ In consequence, stereotypical inferences supported only by the dominant sense are made also from the subordinate use. This explains experimentally documented spatial inferences (*X is in front of Y*) from purely epistemic uses of 'see' ('He saw her point') (Fischer and Engelhardt 2017, 2019,

⁸ The two most relevant principles of activation are that senses get activated more strongly when encountered more frequently than other senses (Giora 2003), and that frequently co-activated stereotype components exchange lateral cross-activation (McRae et al. 2005).

2020), doxastic inferences (*S believes that X is F*) from purely phenomenal uses of appearance verbs ('X appears F to S') (Fischer et al. 2021, Fischer, Engelhardt, and Sytsma 2021), and inferences of stereotypical zombie properties from philosophical uses of 'zombie' (Fischer and Sytsma 2021). Professional academic philosophers proved as susceptible to this bias as laypeople (Fischer, Engelhardt, and Herbelot 2022).

These empirical findings warrant new reconstructions of arguments from illusion and hallucination that develop the problem of perception. These arguments employ appearance- and perception-verbs in a 'phenomenal' or 'phenomenological' sense (Ayer 1956, p. 90, Robinson 1994, pp. 51–53) which lacks many of the implications – factive, epistemic, or doxastic – of ordinary uses of these verbs, as it serves only to describe the character of perceivers' subjective experience (e.g., 'Macbeth sees_{SPHEN} a dagger' means 'Macbeth had an experience like that of seeing a dagger'). The linguistic salience bias hypothesis predicts that even competent thinkers will make contextually inappropriate stereotypical inferences from these uses that lead to fallacies of equivocation.

For example, an influential statement of the argument from hallucination assumes that

(1) Macbeth saw a dagger,

and immediately infers 'but still not a real dagger' (Ayer, 1956 p. 90). Since the hallucinator is meant to have an experience just like that of seeing a physical dagger, this inference is fallacious: In the relevant phenomenal sense, Macbeth is seeing a physical dagger – he merely fails to see such a thing in the dominant visual sense. The linguistic salience bias hypothesis suggests the step is driven by existential or spatial inferences that are supported by the dominant visual sense but are made from the special phenomenal use in (1), anyway.⁹ They lead to implicit conclusions that are defeated by contextual information but are tacitly assumed, even so, in further reasoning:

(2) There was something there that Macbeth saw. But, by assumption:

(3) There was no physical object there for Macbeth to see. By (2) and (3):

(4) There was a non-physical object that Macbeth saw.

Here, the inference from (1) leads to an implicit conclusion (2) that is immediately defeated (by 3) but tacitly assumed in the inference from (1) and (3) to (4) (see Fischer and Engelhardt 2020, pp. 432–433). The second half of the argument then generalises from this case of hallucination to all – qualitatively indistinguishable – cases of visual perception (Ayer 1956, p. 90, McPherson 2013, pp. 15–16, Smith 2002, pp. 196–197).

Such diagnostic reconstructions of arguments from hallucination (Fischer 2019, Fischer and Engelhardt 2020) and from illusion (Fischer et al. 2021, Fischer, Engelhardt, and Sytsma 2021) promise to 'dissolve' the problem of perception by showing that the apparent conflict between belief in the occurrence or possibility of illusions and hallucinations, on the one hand, and folk direct realism, on the other, is only due to fallacious reasoning. The puzzle is an artefact of a verbal reasoning bias.

Such diagnostic accounts have two key elements: reconstructions of specific arguments that expose fallacies in these arguments, and explanations of why even competent thinkers

⁹ On some interpretations of (1), Macbeth hallucinates a purely imaginary dagger; on others, he hallucinates the actual murder weapon he will use (which just is not around at this point), and the following steps (2) and (3) (below) quantify over locally restricted domains, with the second 'there' understood as 'before Macbeth'. In the former case, the argument involves parallel existential and spatial inferences from (1); in the latter case, the argument relies exclusively on spatial inferences (Fischer and Engelhardt 2020).

commit those fallacies, e.g., by reference to a cognitive bias. Both elements are in need of empirical support. To support the proposed reconstruction of the argument from hallucination, for instance, we need to document factive inferences from phenomenal uses of perception-verbs. To support the proposed explanation, we need to provide evidence of linguistic salience bias, more generally. Methods from psycholinguistics allow us to address both tasks.

3.3 Psycholinguistic Methods

We now illustrate how psycholinguistic methods that were initially developed to study language comprehension can be adapted for the new purpose of studying verbal reasoning. To do so, we consider experiments that assessed the linguistic salience bias hypothesis.

Hypotheses about automatic inferences from words can be tested with the psycholinguistic cancellation paradigm (for a review, see Fischer and Engelhardt 2019): Participants read or hear sentences where the word of interest is followed by sequels that are, respectively, consistent and inconsistent with the conclusion of the hypothesized inference. If these inferences are made, the resulting clashes with inconsistent sequels engender comprehension difficulties which require effort to overcome. Increased cognitive effort shows up through pupil dilations (Laeng et al. 2012), longer reading times (Rayner et al. 2004), and signature electrophysiological responses in the brain (Kutas and Federmeier 2011). Experimental philosophers have taken up all these methods: pupillometry (e.g., Fischer and Engelhardt 2017, 2020), reading-time measurements with eye tracking (e.g., Fischer, Engelhardt, and Sytsma 2021), as well as electroencephalography (e.g., Cosentino et al. 2017). We discuss the method that strikes us as the most helpful for the investigation of inferences in philosophical arguments: reading-time measurements with eye-tracking, combined with plausibility ratings.

Readers fixate most words as their eyes move forward (skipping the words easiest to predict from the context) *and* backwards ('regressions' at points of difficulty). Readers immediately construct local interpretations over small numbers of adjacent words. If the task at hand demands it, and only then, they subsequently integrate these local interpretations into more comprehensive interpretations of longer sentences and passages, which take long-distance dependencies into account (Swets et al. 2008; cf. Ferreira and Lowder 2016). In reading, we thus need to recognise words and integrate them into local and more comprehensive interpretations. Difficulties at these different stages manifest themselves in different eye-tracking measures (first-pass reading times, re-reading times, and so forth). Where default inferences from words get cancelled by subsequent context, readers have difficulties integrating information from the initial 'source region' with information from the subsequent cancellation or 'conflict region'. Such difficulties lead to longer re-reading times, for either of these two regions (Rayner et al. 2004, Clifton et al. 2016). Automatic inferences can therefore be studied by examining re-reading times.

Fischer and Engelhardt (2019) used this approach to document inappropriate spatial inferences from epistemic uses of 'see', predicted by their linguistic salience bias hypothesis. Materials manipulate verb, object, and post-verbal context, in a 2×2×2 design (see below, Table 1). Concrete vs. abstract objects (e.g., 'picture' vs. 'problems') invite visual vs. epistemic interpretations of the verb. Post-verbal contexts are amenable to literal spatial and metaphorical temporal interpretation (whereby ahead = in the future; behind = in the past). These facilitate purely metaphorical interpretations of epistemic items (*Joe knows what problems he will have in the future* and *Joe knows what problems he had in the past*). If any spatial inferences from

‘see’ are suppressed from the local interpretation, e.g., of ‘Joe sees the problems’, readers have no trouble winning through to this interpretation. The space–time metaphors in our items give rise to embodied cognition effects (Boroditsky and Ramscar 2002, Bottini et al. 2015) and support spatial reasoning about temporal relations (Casasanto and Boroditsky 2008, Gentner, Imai and Boroditsky 2002). If spatial inferences from ‘see’ cannot be suppressed from the local interpretation of the previous words, they will engage spatial reasoning, and engender integration difficulties in sentences like ‘Joe sees the problems that lie behind him’. Such inferences will hence show up through longer re-reading times for spatially inconsistent than spatially consistent contexts.

	Verb	Object		Context	
<u>Visual</u>					
1.	Sheryl sees	the picture	on the wall	behind her.	(s-inconsistent)
2.	Sheryl sees	the picture	on the wall	facing her.	(s-consistent)
3.	Sheryl is aware of	the picture	on the wall	behind her.	(s-inconsistent)
4.	Sheryl is aware of	the picture	on the wall	facing her.	(s-consistent)
<u>Epistemic:</u>					
1.	Joe sees	the problems	that lie	ahead of him.	(s-consistent)
2.	Joe sees	the problems	that lie	behind him.	(s-inconsistent)
3.	Joe is aware of	the problems	that lie	ahead of him.	(s-consistent)
4.	Joe is aware of	the problems	that lie	behind him.	(s-inconsistent)

Table 1. Example stimuli and regions of interest for eye movement analysis (from Fischer and Engelhardt 2019).

Even when they persist long enough to cause integration difficulties, inappropriate default inferences may be suppressed within one second. To examine their influence of further cognition, Fischer and Engelhardt (2019) therefore combined the reading task with a subsequent plausibility rating task: If spatial inferences persist long enough to influence further judgment and reasoning, they will engender the impression of a conflict in spatially inconsistent items, which will reduce the items’ plausibility. Accordingly, lower plausibility ratings for spatially inconsistent than consistent ‘see’-sentences provide evidence that spatial inferences from ‘see’ influence further cognition, as do lower ratings for spatially inconsistent sentences with ‘see’ than with the contrast verb ‘be aware of’ that does not give rise to persistent spatial inferences. This experimental paradigm has been used to garner evidence for contextually inappropriate stereotypical inferences. It permits to examine, more generally, automatic inferences in verbal reasoning. Findings can provide much-needed empirical foundations for the promising approach of critical ordinary language philosophy, and assess diagnostic analyses proposed for specific philosophical arguments.

4. Constructive Ordinary Language Philosophy: Corpus Analyses

In contrast to the critical project that seeks to dissolve (‘Platonic’) philosophical puzzles, constructive ordinary language philosophy aims to draw positive conclusions about phenomena of philosophical interest. As Hansen (2020, p. 2434) puts it, ‘the constructive project consists of observations about how philosophically significant expressions are ordinarily used and uses those observations to support conclusions about non-linguistic aspects of the world’. Of course, the inferences one draws from how we talk about the world to how the world actually is can be

stronger or weaker – and one can be more or less sceptical about such inferences. Minimally, we can see constructive ordinary language philosophy as seeking to use ‘a sharpened awareness of words to sharpen our perception of, though not as the final arbiter of, the phenomena’ (Austin 1957, p. 8). The hope is that such sharpened perception will result in improved understanding of the non-linguistic phenomena and, in particular, their roles in our lives. This enterprise has an empirical side: empirical methods are required to investigate how we use words, the situations in which we apply them, and the ends to which they are put.

The empirical side of constructive ordinary language philosophy corresponds with what Sytsma and Livengood (2015) term the *descriptive programme* of experimental philosophy. As they put it, practitioners ‘aim to describe how people actually talk about philosophically interesting topics’ (p. 43), hoping to detail ‘the contours of our concepts’ (p. 76). While Sytsma and Livengood focus on how this can be done by using the classic, case-based questionnaire methods, we now consider how constructive ordinary language philosophy can be aided by an expanded set of tools, especially tools from corpus linguistics. As an example, we look at Sytsma and Livengood’s responsibility account of ordinary causal attributions. We now introduce work on causal attributions (Section 4.1) and corpus methods (Section 4.2), and suggest how findings can serve the aims of constructive ordinary language philosophy (Section 4.3).

4.1 Causal Attributions and Injunctive Norms

What is causation? The recent philosophical literature has tended to treat causation as a non-linguistic aspect of the world, and yet has focused on ‘our folk-theoretical notion of “cause”’ (Paul and Hall 2003, p. 2) in examining it, often calling on case judgments about the applicability of this lemma¹⁰ for purposes of giving an analysis of the concept, which is in turn treated as evidence with regard to the nature of causation. Frequently, the cases employed have had a clear normative dimension. For example, Carolina Sartorio (2005) calls on several cases involving life and death decisions to set ‘a constraint on the concept of cause’ and thereby help ‘carve up the concept’ (p. 71). The first is the Assassination Case, which is taken to provide an obvious counterexample to a simple counterfactual account of causation:

Assassin shoots Victim and, as a result, Victim dies. However, Backup is waiting in reserve. Had Assassin not shot, Backup would have, and Victim would still have died (in a very similar way, at around the same time, etc.). (p. 72)

Despite calling on norm-laden cases like this, philosophers have generally assumed that the concept of causation is a purely descriptive concept, such that the normative dimension might be seen as simply adding a bit of colour to the discussion. As Helen Beebe (2004) puts it, ‘no philosopher working within the tradition I’m concerned with here thinks that the *truth* conditions for causal claims contain a moral element’ (p. 293).

Seminal work in experimental philosophy has cast grave doubt on the assumption that the moral aspects of cases can simply be ignored, however, at least insofar as we hope to analyse the folk concept of causation (see Chapter 9). Perhaps the most discussed example, the Pen Case, comes from Knobe and Fraser (2008, p. 443):

¹⁰ A lemma is the dictionary form of a set of words or the word they are indexed by. To illustrate, the concern here is presumably not simply with the word ‘cause’ but would also include ‘caused’, ‘causes’, and ‘causing’.

The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take the pens, but faculty members are supposed to buy their own.

The administrative assistants typically do take the pens. Unfortunately, so do the faculty members. The receptionist has repeatedly e-mailed them reminders that only administrative assistants are allowed to take the pens.

On Monday morning, one of the administrative assistants encounters Professor Smith walking past the receptionist's desk. Both take pens. Later that day, the receptionist needs to take an important message... but she has a problem. There are no pens left on her desk.

In this case, two agents perform the same action, jointly bringing about a bad outcome. The only difference is that one violates an injunctive norm in doing so – in this case doing something that is proscribed by established rules – while the other does not. Despite this, when participants are asked who caused the problem, ratings are very notably higher for the norm-violating agent (Professor Smith) than for the norm-conforming agent (the administrative assistant). And similar findings have been found using a range of different vignettes, using both within- and between-participant designs, and employing a variety of types of questions, including using Likert-scales, binary and multiple-choice answers, rank-ordering, and fill-in-the-blank (e.g., Alicke 1992, 2000, Hitchcock and Knobe 2009, Kominsky et al. 2015, Icard et al. 2017, Kominsky and Phillips 2019, Sytsma and Livengood 2021, Sytsma forthcoming-a).

One key question that the experimental literature has focused on is how we should explain the effect of norms on causal attributions. A number of different types of accounts have been offered, although we'll focus on just two here.¹¹ Perhaps the most prominent are *counterfactual accounts*, such as that offered by Hitchcock and Knobe (2009), building off the work of Knobe and Fraser (2008) discussed above.¹² Hitchcock and Knobe's work can be categorized most cleanly as an example of explanatory experimental philosophy: they note the effect of norms on causal attributions and lay out a cognitive mechanism to explain it. Specifically, they hold that norm violations matter for causal attributions because people arrive at such judgments by considering counterfactuals, but don't do so indiscriminately; rather, they are more likely to consider counterfactuals on which a norm violation is replaced with something more normal. If the outcome doesn't occur on this counterfactual, they then treat the abnormal event as the cause. For instance, in the Pen Case, Professor Smith violates a norm in taking a pen (while the administrative assistant does not). According to Hitchcock and Knobe, this renders Professor Smith's action more salient, with people then being more likely to consider the counterfactual on which Professor Smith does not take a pen. Since the problem doesn't occur on this counterfactual, people then tend to judge that Professor Smith caused the problem. This type of view is in accord with the philosophical consensus, treating the ordinary concept of causation as purely descriptive, while explaining the impact of norms indirectly, holding that norms affect the process by which we apply the descriptive concept.

Sytsma, Livengood, and Rose (2012) offered a rather different type of explanation. Pushing against the philosophical consensus, they suggested that the ordinary concept of causation is not purely descriptive, but in part normative, such that ordinary causal attributions are akin to responsibility or accountability attributions.

¹¹ See also the bias account put forward by Alicke and colleagues (e.g., Alicke 1992, 2000, Alicke, Rose, and Bloom 2011) as well as the pragmatic account put forward by Samland and Waldmann (2016).

¹² This type of account has been developed in a large number of papers, including Halpern and Hitchcock (2015), Kominsky et al. (2015), and Icard et al. (2017), among others.

Projects in experimental philosophy seldom map cleanly onto theoretical divisions between programmes, and the work of Sytsma and colleagues is no exception. Thus, the responsibility account has been pushed in ways that align with evidential experimental philosophy, stressing the divergence between ordinary causal attributions and philosophical assumptions (e.g., Livengood, Sytsma, and Rose 2017, Livengood and Sytsma 2020, Sytsma forthcoming-b). And, in line with explanatory experimental philosophy, the responsibility account is given as an explanation of effects like that shown by Knobe and Fraser (2008) for the Pen Case. Nonetheless, Sytsma and colleagues' primary claim concerns how English speakers typically use the lemma 'cause', and the main goal of their work is to help detail the contours of this usage. In other words, the evidential and explanatory ends follow from a project starting in descriptive experimental philosophy. This is perhaps most clear in the work of Sytsma and colleagues (2019), which goes beyond the case-based questionnaire methods that have most commonly been employed in experimental philosophy of causation by calling on the methods of corpus linguistics.

4.2 Corpus Methods

While questionnaire methods have been the dominant approach in experimental philosophy, they are not the only methods that have been used. One set of methods that has been becoming increasingly popular in recent years are corpus methods, which have often been employed with the goal of supplementing data collected in experimental settings with information about the use of words 'in the wild' so as to provide a consilience of evidence (Bluhm 2016, Ulatowski et al. 2020).¹³

Corpus linguistics aims to collect and analyse pre-existing, 'real world' linguistic data (McEnery and Hardie 2015, O'Keefe and McCarthy 2022). Corpus analysis starts with corpora. These are collections of written or oral texts, typically supplemented with additional information (e.g., where the texts were drawn from, parts of speech, and so on). Corpora are often built with the aim of giving a balanced picture of the target domain, which might be rather general (e.g., contemporary American English) or quite specific (e.g., the utterances of children). Many corpora are available online, often with extensive search features that, among other things, allow users to easily compare how often different words occur and the contexts in which they are used. Perhaps the most common corpus in recent work in experimental philosophy is the Corpus of Contemporary American English (COCA).¹⁴ This is a contemporary, balanced corpus, comprised of some 1 billion words drawn from a range of sources (including magazines and newspapers, blogs and webpages, books and academic texts, transcripts and subtitles). And COCA is the corpus used by Sytsma and colleagues (2019).

Their main goals were to determine whether the use of causal attributions tends to be sensitive to normative information outside of experimental contexts and whether they are similar to responsibility attributions, as predicted by the responsibility account. To do this, they used the COCA in two different ways: They began by using the online search features to compare contexts for key phrases of interest, then confirmed their findings by applying computational methods to the full text of the corpus.

¹³ While the use of corpus methods in philosophy dates back to at least the 1970's (e.g., Meunier et al. 1976, McKinnon 1970), its use has grown notably in the past five years. For an extensive, if incomplete, set of references, see Sytsma (forthcoming-c, fn. 3).

¹⁴ <https://www.english-corpora.org/coca/>

As a proxy for the presence of normative information in the texts, Sytsma and colleagues looked at what events people tend to attribute causation for, noting that people are generally more focused on blame than praise (Prinz 2007). As such, they expected that if causal attributions are used in a normative way, we would expect them to be disproportionately directed toward bad outcomes. By contrast, if the ordinary concept of causation is purely descriptive, we would have no a priori reason to expect it to be used in contexts with any particular valence. To test this, Sytsma and colleagues compiled the nouns that occur most frequently after the phrases standardly used to elicit causal attributions (‘caused the’) and responsibility attributions (‘responsible for the’) in the experimental literature, as well as a range of possible synonyms, using the non-academic portions of the COCA. This resulted in a list of 260 distinct items. They then had independent coders judge the valence of each item.

As expected, Sytsma and colleagues found that ‘caused the’ was disproportionately used to describe items that the coders judged to be negative. To illustrate, the five most common nouns were ‘death’, ‘accident’, ‘crash’, ‘problem’, and ‘explosion’. And a similar result held for responsibility attributions, but not for most of the potential synonyms tested, such as ‘created the’ or ‘produced the’, which were mainly used in a neutral fashion. Based on these results, as well as further analyses aimed at testing potential confounds, Sytsma and colleagues (2019) conclude that ‘the data suggests that the notion of causation is in itself inherently normative’ (p. 223).

To further test this finding, Sytsma and colleagues turned to more computational methods, constructing distributional semantic models (DSMs) from the full text of the non-academic portions of the COCA. While the details can be daunting, the basic idea behind this approach is that ‘you shall know a word by the company it keeps’ (Firth 1957, p. 11).¹⁵ To do this, DSMs look at the contexts in which terms occur across the corpus, typically representing each term as a geometric vector in a high-dimensional space. Closeness in the space is interpreted in terms of relatedness of meaning, and standardly measured by taking the cosine of the vectors. In their final DSM, Sytsma and colleagues used the *word2vec* algorithm to build a high-performing semantic space¹⁶ with the key phrases from above (‘caused the’, ‘responsible for the’) treated as individual terms. As predicted, they found that these were quite close together in the semantic space, indicating that they are used in very similar ways. In fact, each of them was one of the five nearest terms to the other in the space. Further, other close terms suggested, again, that each is used in predominately negative contexts. These findings lend further strength to the more qualitative analysis, as well as prior experimental results, supporting the key contention of the responsibility account – that causal attributions are typically used in a normative fashion.

4.3 From Descriptive to Constructive

The responsibility account suggests that the folk concept of causation is normative. In ‘descriptive experimental philosophy’ (Sytsma and Livengood 2015), such a descriptive conclusion is treated as an end in itself. But it can also form the basis for inferences about non-

¹⁵ For more detailed discussion of DSMs, see Baroni et al. (2014a), Erk (2012), as well as Turney and Pantel (2010). For further examples illustrating different philosophical applications, see Fischer, Engelhardt, and Herbelot (2015; 2022), Fischer and Sytsma (2021), Sytsma and Snater (forthcoming), Zahorec et al. (forthcoming).

¹⁶ As measured by the MEN benchmark (Bruni et al. 2013), for which it produced a Spearman’s ρ of 0.80, which is on par with the best performing models in Baroni et al. (2014b).

linguistic phenomena. In the case at hand, the sharpened perception of the ordinary use of ‘cause’ plausibly suggests a shift in how philosophers should think about the phenomenon of causation, resituating it from the non-social to the social realm: Rather than see the normative aspect of our causal attributions as reflecting a mistake or a bias that threatens to distract us from the a more fundamental descriptive relation, we can recalibrate our thinking to recognize the centrality of normative concerns to our daily lives. Thus, Sytsma (2021) suggests that the responsibility account fits neatly with a general view of how humans tend to think about the world—that ‘we are moralizing creatures through and through’ (Knobe 2010, p. 328), tending to focus not just on how things came about, but also on how to apportion praise and blame for what has occurred. This is exactly the kind of conclusion constructive ordinary language philosophy aims for: a conclusion that is drawn from an examination of the ordinary uses of words and that helps to situate our thinking about non-linguistic phenomena.

5. Conclusion: Experimental Philosophy and the Kantian Ambition

This chapter considered how experimental philosophy can help realize the Kantian ambition of placing philosophy – or notable parts of it – on the secure path of a science. Among other things, experimental philosophy uses scientific methods to investigate how people think and speak about phenomena of philosophical interest. To discern how such an investigation can help realize the Kantian ambition, we need to understand how an improved understanding of how people think and speak about phenomena can help address philosophical questions and problems about these phenomena.

We have suggested that two main strands of experimental philosophy potentially face a dilemma: the critic can contend that evidential experimental philosophy at best makes indirect contributions to philosophy, while explanatory experimental philosophy reduces to mere psychology. To identify ways in which experimental philosophy can play a more substantive role and contribute more directly to addressing characteristically philosophical problems, we turned to its most direct historical precursor, ordinary language philosophy.

Constructive ordinary language philosophy seeks to advance from claims about folk concepts to a better understanding of the non-linguistic phenomena these concepts are used to conceptualise, as well as the role these phenomena play in our lives. We have noted that descriptive experimental philosophy offers a base for this practice, aiming to detail the contours of our concepts. And we have suggested that methods from corpus linguistics can play an important role in the practice, supplementing more common questionnaire methods. By calling on a diversity of methods, we can gain insight into our language, and with it our underlying concepts, setting the stage for constructive inferences about the non-linguistic world.

Critical ordinary language philosophy assumes that many philosophical puzzles – ‘Platonic puzzles’, as we called them – result from fallacious reasoning about the phenomena of interest. The approach then seeks to resolve such puzzles by exposing fallacies in the relevant reasoning. We considered an experimental implementation of the approach that addresses the ‘problem of perception’ that is engendered by arguments from illusion and hallucination. The implementation develops and assesses explanations of pertinent fallacies, and documents the particular fallacious inferences that are hypothesised to drive the targeted puzzle-generating arguments. Diagnostic analyses that expose fallacies in the early stages of all arguments that contribute to the relevant puzzle would then ‘dissolve’ this puzzle (e.g., the problem of perception).

A potentially complementary approach builds on findings from experimental philosophy that suggest persistent inter- or even intrapersonal conflicts between beliefs. The negative program's investigation of intuitional sensitivity prompted still ongoing debates about the extent and relevance of such sensitivity (Knobe 2021, Stich and Machery 2022). Regardless of how these debates pan out, however, this work has revealed a considerable amount of interpersonal conflict (for a review, see Knobe 2021): Both where intuitions of interest proved sensitive to the epistemically irrelevant parameters investigated and where they proved stable, considerable minorities disagreed with the majority judgments, and were typically too large to be easily explained away as performance errors. Indeed, in some cases, people across cultures appear to be almost evenly split between competing judgments (e.g., compatibilism vs. incompatibilism about free will; see Hannikainen et al. 2019). In addition, sensitivity to presentational factors may sometimes be indicative of intrapersonal conflicts: People may feel torn between opposing judgments, and even minor presentational factors, such as small differences in wording, may sway them in different directions.

Such intrapersonal conflicts may be the ultimate source of many philosophical puzzles and problems. 'Fragmentation accounts' of belief storage suggest that people's beliefs are often 'fragmented' and conflicted (Bendaña and Mandelbaum 2021, Leiser 2001): Different cognitive processes, operating under different conditions, generate conflicting beliefs, which are never systematically screened for coherence; they are stored at different locations in long-term memory, in different 'belief fragments', which are internally coherent, but may conflict with one another. As a result, different, and conflicting conceptions of the same phenomenon may be held not only by members of different communities (as findings from Berniūnas et al. 2021 suggest for free will) and different members of the same community (as findings from Latham, Miller, and Norton 2021 suggest for time), but even by the same individual (as findings from Fischer, Allen, and Engelhardt n.d. suggest for vision).

Philosophical puzzles may thus arise from direct conflicts between different belief fragments. For instance, the problem of perception may be exacerbated by arguments from illusion and hallucination, but ultimately arise from a conflict between equally pre-scientific beliefs about vision that are consistent with direct realism and indirect realism, respectively, and are frequently held by the same individuals (Fischer, Allen, and Engelhardt n.d.). In this case, debunking explanations of one of the two parties to the conflict could help 'dissolve' the puzzle. For instance, a grossly simplified implicit model of attention that is used to track others' and one's own focus of visual attention (Webb and Graziano 2015) could account for indirect realist beliefs about vision. The resulting explanation would debunk these beliefs and thus contribute to 'dissolving' the problem of perception.

Alternatively, empirical insight into how fragmented and conflicted folk belief are may have us regard conflicts between folk beliefs as unavoidable parts of the human condition. Due to their incoherence, we would then simply set aside folk beliefs in efforts to explain phenomena and would accept as unavoidable any cognitive dissonance that arises from conflicts between different belief fragments. This attitude would lead us to stop bothering about philosophical puzzles that arise from such conflicts. Either way, be it through diagnostic analyses, debunking explanations, or meta-cognitive insight into the workings and extent of belief fragmentation, empirical insights of a kind to which experimental philosophy aspires can make key contributions to overcoming longstanding puzzles and problems of a distinctively philosophical nature.

About the Authors

Eugen Fischer (DPhil Oxford) is Reader in Philosophy at the University of East Anglia. He is the author of *Linguistic Creativity* (2000) and *Philosophical Delusion and its Therapy* (2011). He has co-edited several volumes including *Experimental Philosophy, Rationalism, and Naturalism* (2015) and *Methodological Advances in Experimental Philosophy* (2019).

Justin Sytsma (PhD University of Pittsburgh) is Associate Professor in the Philosophy Programme at Victoria University of Wellington. He is co-author of *The Theory and Practice of Experimental Philosophy* (2015), co-editor of *A Companion to Experimental Philosophy* (2016), and editor of *Advances in Experimental Philosophy of Mind* (2014).

Bibliography

- Alexander, Joshua, Chad Gonnerman, and John Waterman (2014). Salience and epistemic egocentrism: An empirical study. In J. Beebe (Ed.), *Advances in Experimental Epistemology* (pp. 97-118). London: Bloomsbury.
- Alicke, Mark (1992). Culpable causation, *Journal of Personality and Social Psychology*, 63, 368–378.
- Alicke, Mark (2000). Culpable Control and the Psychology of Blame, *Psychological Bulletin*, 126 (4), 556–574.
- Alicke, Mark (2008). Blaming badly. *Journal of Cognition and Culture*, 8, 179-186.
- Alicke, Mark, David Rose, and Dori Bloom (2011). Causation, Norm Violation and Culpable Control, *Journal of Philosophy*, 108, 670-696.
- Allen, Keith (2016). *A Naïve Realist Theory of Colour*. Oxford: OUP.
- Arico, Adam, Brian Fiala, Robert Goldberg, and Shaun Nichols (2011). The folk psychology of consciousness. *Mind and Language*, 26, 327–352.
- Austin, J.L. (1946). Other minds. *Proceedings of the Aristotelian Society*, 20, 148-197.
- Austin, J.L. (1957). A plea for excuses. *Proceedings of the Aristotelian Society*, 57, 1–30.
- Austin, J.L. (1962). *Sense and Sensibilia*. Oxford: OUP.
- Ayer, A.J. (1956). *The Problem of Knowledge*. London: Penguin.
- Bargh, John, Kay Schwader, Sarah Hailey, Rebecca Dyer, and Erica Boothby (2012). Automaticity in social-cognitive processes. *Trends in Cognitive Sciences*, 16, 593-605.
- Baroni, Marco, Raffaella Bernardi, and Roberto Zamparelli (2014a). Frege in space, *Linguistic Issues in Language Technology*, 9, 241–346.
- Baroni, Marco, Georgiana Dinu, and German Kruszewski (2014b). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors, *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 238–247.
- Bealer, George (1996). On the possibility of philosophical knowledge. *Philosophical Perspectives*, 10, 1–34.
- Bealer, George (2000). A theory of the a priori. *Pacific Philosophical Quarterly*, 81, 1–30.
- Beebe, Helen (2004). Causing and Nothingness. In Collins, Hall, and Paul (eds.), *Causation and Counterfactuals*, MIT Press.
- Beisbart, Claus and Kevin Reuter (2021). What is the folk concept of life? *Australasian Journal of Philosophy*, <https://doi.org/10.1080/00048402.2021.1995449>

- Bendaña, Joseph and Eric Mandelbaum (2021). The fragmentation of belief. In Cristina Borgoni, Dirk Kindermann, and Andrea Onofri (Eds.), *The Fragmented Mind*. Oxford: OUP.
- Berniūnas, Renatas, Audrius Beinorius, Vilius Dranseika, Vytis Silius, and Paulius Rimkevičius (2021). The weirdness of belief in free will. *Consciousness and Cognition*, 87, 103054
- Bicknell, Klinton, Jeffrey Elman, Mary Hare, Ken McRae, and Marta Kutas (2010). Effects of event knowledge in processing verbal arguments. *Journal of Memory and Language*, 63, 489–505.
- Bluhm, Roland (2016). Corpus Analysis in Philosophy. In Martin Hinton (Ed.), *Evidence, Experiment, and Argument in Linguistics and the Philosophy of Language*, Berlin: Peter Lang, pp. 91–109.
- Boroditsky, Lera and Michael Ramscar (2002). The roles of body and mind in abstract thought. *Psychological Science*, 13, 185-188.
- Bottini, Roberto, Davide Crepaldi, Daniel Casasanto, Virgine Crollen, and Olivier Collignon (2015). Space and time in the sighted and blind. *Cognition*, 141, 67-72.
- Brocher, Andreas, Jean-Pierre Koenig, Gail Mauner, and Stephani Foraker (2018). About sharing and commitment: the retrieval of biased and balanced irregular polysemes. *Language, Cognition and Neuroscience*, 33:4, 443-466,
- Bruni, Elia, Nam Khahn Tran, and Marco Baroni (2013). Multimodal Distributional Semantics, *Journal of Artificial Intelligence Research*, 49, 1–47.
- Buckwalter, Wesley (2010). Knowledge isn't closed on Saturdays. *Review of Philosophy and Psychology*, 1, 395–406.
- Butterfuss, Reese and Panayiota Kendeou (2018). The role of executive functions in reading comprehension. *Educational Psychology Review*, 30, 801–826
- Callender, Craig (2017). *What Makes Time Special?* Oxford: OUP.
- Casasanto, Daniel and Lera Boroditsky (2008). Time in the mind: Using space to think about time. *Cognition*, 106, 579-593.
- Chalmers, David (1996). *The Conscious Mind*. New York: OUP
- Chalmers, David and Frank Jackson (2001). Conceptual analysis and reductive explanation. *Philosophical Review*, 110, 315-360.
- Clifton, Charles, Fernanda Ferreira, John Henderson, Albrecht Inhoff, Simon Livesedge, Erik Reichle, E.D., and Elizabeth Schotter (2016). Eye movements in reading and information processing: Keith Rayner's 40 year legacy, *Journal of Memory and Language*, 86, 1–19.
- Cohen, Jonathan and Shuan Nichols (2010). Colours, colour relationalism and the deliverances of introspection, *Analysis*, 70, 218-228.
- Collins, John (2015). Naturalism without metaphysics. In Eugen Fischer and John Collins (Eds.), *Experimental philosophy, Rationalism, and Naturalism* (pp. 85-109). London: Routledge
- Cosentino, Erica, Giosue Baggio, Jarmo Kontinen, and Markus Werning (2017). The time-course of sentence meaning composition. N400 effects of the interaction between context-induced and lexically stored affordances, *Frontiers in Psychology*, 8 (818).
<https://doi.org/10.3389/fpsyg.2017.00813>

- Cova, Florian (2016). The folk concept of intentional action: Empirical approaches. In Justin Sytsma and Wesley Buckwalter (Eds.), *A Companion to Experimental Philosophy* (pp. 121–141). Malden: Wiley Blackwell.
- Crane, Tim and Craig French (2021). The problem of perception. In Edward Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/fall2021/entries/perception-problem>
- Daly, Christopher (2010). *An Introduction to Philosophical Methods*. Broadview
- De Caro, Mario and David Macarthur (ed.) 2008. *Naturalism in Question*. Cambridge: HUP.
- DeRose, Keith (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, 52, 913–29.
- Erk, Katrin (2012). Vector space models of word meaning and phrase meaning: A survey, *Language and Linguistics Compass*, 6 (10), 635–653.
- Engelhardt, Paul and Fernanda Ferreira (2016). Reaching sentence and reference meaning. In Pia Knoeferle, Pirita Pykkonen-Klauck, and Matthew Crocker (Eds.), *Visually Situated Language Comprehension* (pp. 127-150). John Benjamins
- Faust, Mark and Morton Gernsbacher (1996). Cerebral mechanisms for suppression of inappropriate information during sentence comprehension, *Brain and Language*, 53, 234-259.
- Ferreira, Fernanda and Matthew Lowder (2016). Prediction, information structure, and good-enough language processing, *Psychology of Learning and Motivation*, 65, 217–247.
- Firth, J. R. (1957), A synopsis of linguistic theory 1930–55. In *Studies in Linguistic Analysis*, Oxford: Blackwell, pp. 1–32.
- Fischer, Eugen (2011). *Philosophical Delusion and its Therapy*. New York: Routledge.
- Fischer, Eugen (2019). Linguistic Legislation and Psycholinguistic Experiments: Redeveloping Waismann’s Approach. In D. Makovec and S. Shapiro (Eds.), *Friedrich Waismann*. Palgrave Macmillan
- Fischer, Eugen (2020). Conceptual control. On the feasibility of conceptual engineering, *Inquiry*, 1-29. <http://dx.doi.org/10.1080/0020174X.2020.1773309>
- Fischer, Eugen, Keith Allen, and Paul Engelhardt. *Fragmented and conflicted: Folk beliefs about vision*. Unpublished ms. University of East Anglia.
- Fischer, Eugen and John Collins (2015). Rationalism and naturalism in the age of experimental philosophy. In their (Eds.), *Experimental philosophy, Rationalism, and Naturalism* (pp.1-33). London: Routledge
- Fischer, Eugen and Mark Curtis (Eds.) (2019). *Methodological Advances in Experimental Philosophy*. London: Bloomsbury.
- Fischer, Eugen and Paul Engelhardt (2017). Stereotypical inferences: Philosophical relevance and psycholinguistic toolkit, *Ratio*, 30, 411–442.
- Fischer, Eugen and Paul Engelhardt (2019). Eyes as windows to minds: Psycholinguistics for experimental philosophy. In Eugen Fischer and Mark Curtis (eds.), *Methodological Advances in Experimental Philosophy* (pp.43-100). Bloomsbury.
- Fischer, Eugen and Paul Engelhardt (2020). Lingering stereotypes: Salience bias in philosophical argument, *Mind and Language*, 35, 415-439.
- Fischer, Eugen, Paul Engelhardt, and Aurelie Herbelot (2015). Intuitions and Illusions: From Experiment and Explanation to Assessment. In Eugen Fischer and John Collins (Eds.), *Experimental Philosophy, Rationalism and Naturalism* (pp.259-292). London: Routledge.

- Fischer, Eugen, Paul Engelhardt, and Aurelie Herbelot (2022). Philosophers' linguistic expertise: a psycholinguistic approach to the expertise objection against experimental philosophy, *Synthese*, 200, 1–33. <https://doi.org/10.1007/s11229-022-03487-3>
- Fischer, Eugen, Paul Engelhardt, Joachim Horvath, J., and Hiroshi Ohtani (2021). Experimental ordinary language philosophy: A cross-linguistic study of defeasible default inferences, *Synthese*, 198, 1029–1070.
- Fischer, Eugen, Paul Engelhardt, and Justin Sytsma (2021). Inappropriate stereotypical inferences? An adversarial collaboration in experimental ordinary language philosophy, *Synthese*, 198, 10127–10168.
- Fischer, Eugen and Justin Sytsma (2021). Zombie intuitions, *Cognition*, 215, e104807.
- Foot, Philippa (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 5–15.
- Gawronski, Bertram and Fritz Strack (Eds.) (2012). *Cognitive Consistency: A Fundamental Principle in Social Cognition*. New York: Guilford Press.
- Gentner, Dedre, Mutsumi Imai, and Lera Boroditsky (2002). As time goes by: Evidence for two systems in processing space time metaphors, *Language and Cognitive Processes*, 17, 537-565.
- Gerken, Mikkel, Chad Gonnerman, Joshua Alexander, and John Waterman (2020). Salient Alternatives in Perspective, *Australasian Journal of Philosophy*, 98, 792–810.
- Giora, Rachel (2003). *On Our Mind. Salience, Context, and Figurative Language*. Oxford: OUP.
- Goldman, Alvin (2007). Philosophical Intuitions: Their target, their source, and their epistemic status, *Grazer Philosophische Studien*, 74, 1–26.
- Halpern, Joseph and Christopher Hitchcock (2015). Graded Causation and Defaults, *British Journal for the Philosophy of Science*, 66, 413–457.
- Hannikainen, Ivar, Edouard Machery, David Rose, Stephen Stich, et al. (2019). For whom does determinism undermine moral responsibility? Surveying the conditions for free will across cultures, *Frontiers in Psychology* 10, 2428.
- Hansen, Nat (2020). 'Nobody would really talk that way!': the critical project in contemporary ordinary language philosophy, *Synthese*, 197, 2433-2464.
- Hansen, Nat and Emmanuel Chemla (2015). Linguistic experiments and ordinary language philosophy, *Ratio*, 28, 422–445.
- Hansen, Nat and Emmanuel Chemla (2017). Color adjectives, standards, and thresholds: an experimental investigation, *Linguistics and Philosophy*, 40, 239–278.
- Haug, Matthew (ed.) 2014. *Philosophical Methodology: The Armchair or the Laboratory?* London: Routledge
- Henderson, Leah, (2020). The problem of induction. In Edward Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/spr2020/entries/induction-problem>
- Hitchcock, Christopher and Joshua Knobe (2009). Cause and Norm, *The Journal of Philosophy*, 106, 587–612.
- Horvath, Joachim and Alex Wiegmann (2016). Intuitive expertise and intuitions about knowledge. *Philosophical Studies*, 10, 2701–2726.
- Horvath, Joachim and Alex Wiegmann (2021). Intuitive expertise in moral judgements. *Australasian Journal of Philosophy*. [10.1080/00048402.2021.1890162](https://doi.org/10.1080/00048402.2021.1890162)

- Icard, Thomas, Jonathan Kominsky, and Joshua Knobe (2017). Normality and Actual Causal Strength, *Cognition*, 161, 80–93.
- Jackson, Frank (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. New York: OUP
- Johnston, Mark (1992). How to speak of the colors, *Philosophical Studies*, 68, 221–263.
- Kant, Immanuel (1787/2003). *Critique of Pure Reason*, 2nd ed., trans. N. Kemp Smith. Basingstoke: Palgrave.
- Kintsch, Walter (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163–182.
- Knobe, Joshua (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, 16(2), 309–324.
- Knobe, Joshua (2010). Person as Scientist, Person as Moralist, *Behavioral and Brain Sciences*, 33, 315–365.
- Knobe, Joshua (2016). Experimental philosophy is cognitive science. In Justin Sytsma and Wesley Buckwalter (Eds.), *A Companion to Experimental Philosophy* (pp. 37–52). Wiley.
- Knobe, Joshua (2021). Philosophical intuitions are surprisingly robust across demographic differences. *Filozofia Nauki*, 29 (2), 11–76.
- Knobe, Joshua and Ben Fraser (2008). Causal judgments and moral judgment: Two experiments. In Walter Sinnott-Armstrong (ed.), *Moral Psychology, Volume 2: The Cognitive Science of Morality*, pp. 441–447, Cambridge: MIT Press.
- Kominsky, Jonathan and Jonathan Phillips (2019). Immoral Professors and Malfunctioning Tools: Counterfactual Relevance Accounts Explain the Effect of Norm Violations on Causal Selection, *Cognitive Science*, 43 (11), e12792.
- Kominsky, Jonathan, Jonathan Phillips, Tobias Gerstenberg, David Lagnado, and Joshua Knobe (2015). Causal superseding, *Cognition*, 137: 196–209.
- Kruglanski, Arie, Katarzyna Jasko, Maxim Milyavsky, Marina Chernikova, David Webber, Antonio Pierro, and Daniela di Santo (2018). Cognitive consistency theory in social psychology: A paradigm reconsidered, *Psychological Inquiry*, 29(2), 45–59.
- Kutas, Marta and Kara Federmeier (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP), *Annual Review of Psychology*, 62, 621–647.
- Laeng, Bruno, Sylvain Sirois, and Gustaf Gredebäck (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, 7, 18–27.
- Latham, Andrew, Kristie Miller, and James Norton (2021). Is our naïve theory of time dynamical? *Synthese*, 198, 4251–4271.
- Leiser, David (2001). Scattered naïve theories: Why the human mind is isomorphic to internet web, *New Ideas in Psychology*, 19, 175–202.
- Levin, Janet (2005). The evidential status of philosophical intuition. *Philosophical Studies*, 121, 193–224.
- Levinson, Stephen (2000). *Presumptive meanings. The theory of generalized conversational implicature*. MIT Press.
- Lewis, David (1972). Psychophysical and theoretical identifications, *Australasian Journal of Philosophy*, 50, 249–258.
- Liu, Michelle (2020). The intuitive invalidity of the pain-in-mouth argument, *Analysis*, 80 (3), 463–474.

- Livengood, Jonathan and Justin Sytsma (2020). Actual Causation and Compositionality, *Philosophy of Science*, 87 (1), 43–69.
- Livengood, Jonathan, Justin Sytsma, and David Rose (2017). *Following the FAD: Folk Attributions and Theories of Actual Causation*, *Review of Philosophy and Psychology*, 8 (2), 274–294.
- Ludwig, Kurt (2007). The epistemology of thought experiments: first person versus third person approaches, *Midwest Studies on Philosophy*, 31, 128–159.
- Machery, Edouard (2017). *Philosophy within its Proper Bounds*. Oxford: OUP.
- Machery, Edouard (2021). A new challenge to conceptual engineering, *Inquiry*, 1–24. <https://doi.org/10.1080/0020174X.2021.1967190>
- Macpherson, Fiona (2013). The philosophy and psychology of hallucination. In Fiona Macpherson and Dimitris Platchias (Eds.), *Hallucination: Philosophy and psychology* (pp. 1-38). MIT Press.
- Mallon, Ron (2016). Experimental philosophy. In Herman Cappelen, Tamar Szabo Gendler, and John Hawthorne (Eds.), *Oxford Handbook of Philosophical Methodology* (pp. 410-433). Oxford: OUP.
- Matsuki, Kazunaga, Tracy Chow, Mary Hare, Jeffrey Elman, Christoph Scheepers, and Ken McRae (2011). Event-based plausibility immediately influences on-line language comprehension, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37, 913–934.
- McEnery, Tony and Andrew Hardie (2015). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- McKinnon, Alastair (1970). *The Kierkegaard Indices*, Leiden: Brill.
- McRae, Ken, Mary Hare, Jeffrey Elman, and Todd Ferretti (2005). A basis for generating expectancies for verbs from nouns, *Memory and Cognition*, 33, 1174-1184.
- Metusalem, Ross, Marta Kutas, Thomas Urbach, Mary Hare, Ken McRae, and Jeffrey Elman (2012). Generalized event knowledge activation during online sentence comprehension, *Journal of Memory and Language*, 66, 545–567.
- Meunier, Jean-Guy, Stanislas Rolland, and Francois Daoust (1976). A system for text and content analysis, *Computers and the Humanities*, 10 (5), 281–86.
- Murphy, Taylor (2014). Experimental philosophy 1935-1965. In J. Knobe, T. Lombrozo, and S. Nichols (eds.), *Oxford Studies in Experimental Philosophy: Vol. 1* (pp. 325-368). Oxford: OUP.
- Nadelhoffer, Thomas (2006). Bad acts, blameworthy agents, and intentional actions: Some Problems for juror impartiality, *Philosophical Explorations*, 9, 203-219.
- Nadelhoffer, Thomas, David Rose, Wesley Buckwalter, and Shaun Nichols (2020). Natural compatibilism, indeterminism, and intrusive metaphysics, *Cognitive Science*, 44(8), e12873.
- Naess, Arne (1956). Logical equivalence, intentional isomorphism and synonymity as studied by questionnaires, *Synthese*, 10, 471–479.
- Naess, Arne (1961). A study of ‘or’, *Synthese*, 13, 49–60.
- Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer, and Jason Turner (2004). The phenomenology of free will, *Journal of Consciousness Studies* 11 (7-8), 162–179.
- Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer, and Jason Turner (2005). Surveying freedom: Folk intuitions about free will and moral responsibility, *Philosophical Psychology*, 18 (5), 561–584.

- Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer, and Jason Turner (2006). Is incompatibilism intuitive? *Philosophy and Phenomenological Research*, 73, 28–53.
- Newman, George, Paul Bloom, and Joshua Knobe (2014). Value judgments and the true self, *Personality and Social Psychology Bulletin*, 40, 203–216.
- Nichols, Shaun and Joshua Knobe (2007). Moral responsibility and determinism: the cognitive science of folk intuitions. *Noûs*, 41, 663–685.
- O’Connor, Timothy and Christopher Franklin (2021). Free will. In Edward Zalta (Ed.), *Stanford Encyclopedia of Philosophy*, Spring 2021.
<https://plato.stanford.edu/archives/spr2021/entries/freewill/>
- O’Keefe, Anne and Michael McCarthy (2022). *The Routledge Handbook of Corpus Linguistics, 2nd Edition*. London, New York: Routledge
- Oden, Gregg and James Spira (1983). Influence of context on the activation and selection of ambiguous word senses, *Quarterly Journal of Experimental Psychology*, 35A, 51–64.
- Paul, L.A. and Ned Hall (2003). *Causation: A User’s Guide*. Oxford: Oxford University Press.
- Plato (n.d./2004). *Theaetetus*. Transl. J. Sachs. Newburyport: Focus
- Plunkett, David and Herman Cappelen (2020). A guided tour of conceptual engineering and conceptual ethics. In Harman Cappelen, David Plunkett, and Alexis Burgess (Eds.), *Conceptual Engineering and Conceptual Ethics*. Oxford: OUP.
- Powell, Derek, Zachary Horne, and N. Ángel Pinillos (2014). Semantic integration as a method for investigating concepts. In James Beebe (Ed.), *Advances in Experimental Epistemology*. London: Bloomsbury
- Price, Huw (2011). The flow of time. In Craig Callender (Ed.), *The Oxford Handbook of Philosophy of Time* (pp.276–311). Oxford: OUP.
- Prinz, Jesse (2007), *The Emotional Construction of Morals*. Oxford: OUP.
- Prinz, Jesse (2008). Empirical Philosophy and Experimental Philosophy. In Joshua Knobe and Shaun Nichols (eds.), *Experimental Philosophy*, pp. 189-208, Oxford: OUP.
- Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press
- Rayner, Keith, Tessa Warren, Barbara Juhasz, and Simon Liversedge (2004). The effect of plausibility on eye movements in reading, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1290–1301.
- Reuter, Kevin and Justin Sytsma (2020). Unfelt Pain, *Synthese*, 197, 1777–1801.
- Robb, David and John Heil (2021). Mental causation. In Edward Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/spr2021/entries/mental-causation>
- Roberts, Pendaran, Keith Allen, and Kelly Ann Schmidtke (2018). Folk intuitions about the causal theory of perception, *Ergo*, 3. <https://doi.org/10.3998/ergo.12405314.0003.028>
- Roberts, Pendaran, Keith Allen, and Kelly Ann Schmidtke (2021). Reflective intuitions about the causal theory of perception across sensory modalities, *Review of Philosophy and Psychology*, 12, 257–277.
- Roberts, Pendaran and Kelly Ann Schmidtke (2019). Folk core beliefs about color, *Review of Philosophy and Psychology*, 10, 849-869.
- Robinson, Howard (1994). *Perception*. London: Routledge
- Ryle, Gilbert (1954). *Dilemmas*. Cambridge: CUP.
- Salomons, Tim, Richard Harrison, Nat Hansen, James Stazicker, Astrid Sorensen, Paula Thomas, and Emma Borg (2021). Is pain ‘all in your mind’? Examining the general

- public's views of pain, *Review of Philosophy and Psychology*, <https://doi.org/10.1007/s13164-021-00553-6>
- Samland, Jana and Michael R. Waldmann (2016). How prescriptive norms influence causal inferences, *Cognition*, 156, 164–176.
- Sartorio, Carolina (2005). Causes as Difference-Makers, *Philosophical Studies*, 123, 71-96.
- Schaffer, Jonathan and Joshua Knobe (2012). Contrastive knowledge surveyed. *Noûs*, 675–708.
- Schroeder, Severin (2006). *Wittgenstein: The Way out of the Fly Bottle*. Polity
- Schwenkler, John and Assaf Weksler (2019). Are perspectival shapes seen or imagined? An experimental approach *Phenomenology and the Cognitive Sciences*, 18, 855–877.
- Smith, A.D. (2002). *The Problem of Perception*. Cambridge, Mass: Harvard UP
- Sripada, Chandra (2010). The Deep Self Model and asymmetries in folk judgments about intentional action, *Philosophical Studies*, 151, 159–176.
- Sripada, Chandra (2012). What makes a manipulated agent unfree? *Philosophy and Phenomenological Research*, 85, 563-593.
- Starmans, Christina and Ori Friedman (2020), Expert or esoteric? Philosophers attribute knowledge differently than all other academics, *Cognitive Science*, 44: e12850.
- Stich, Stephen and Edouard Machery (2022). Demographic Differences in Philosophical Intuition: a Reply to Joshua Knobe, *Review of Philosophy and Psychology*, 1–34. <https://doi.org/10.1007/s13164-021-00609-7>
- Stich, Stephen and Kevin Tobia (2016). Experimental philosophy and the philosophical tradition. In Justin Sytsma and Wesley Buckwalter (eds.), *Blackwell Companion to Experimental Philosophy* (pp.5-21). Malden: Wiley Blackwell
- Strawson, P.F. (1959). *Individuals*. London: Methuen.
- Swets, Benjamin, Timothy Desmet, Charles Clifton, and Fernanda Ferreira (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading, *Memory and Cognition*, 36, 201–216.
- Sytsma, Justin (2010). Dennett's Theory of the Folk Theory of Consciousness, *Journal of Consciousness Studies*, 17 (3-4), 107–130.
- Sytsma, Justin (2021). Causation, Responsibility, and Typicality, *Review of Philosophy and Psychology*, 12, 699–719.
- Sytsma, Justin (forthcoming-a). The Responsibility Account. In P. Willemsen and A. Wiegmann (Eds.), *Advances in Experimental Philosophy of Causation*, Bloomsbury.
- Sytsma, Justin (forthcoming-b). *Crossed Wires: Blaming Artifacts for Bad Outcomes*, The Journal of Philosophy.
- Sytsma, Justin (forthcoming-c). *Ordinary Meaning and Consilience of Evidence*. In S. Magen and K. Prochownik (Eds.), *Advances in Experimental Philosophy of Law*, Bloomsbury.
- Sytsma, Justin, Roland Bluhm, Pascale Willemsen, and Kevin Reuter (2019). Causal Attributions and Corpus Analysis. In E. Fischer and M. Curtis (Eds.), *Methodological Advances in Experimental Philosophy*, Bloomsbury, 209-238.
- Sytsma, Justin and Wesley Buckwalter (Eds.) (2016). *A Companion to Experimental Philosophy*. Wiley.
- Sytsma, Justin and Jonathan Livengood (2015). *The Theory and Practice of Experimental Philosophy*. Broadview.
- Sytsma, Justin and Jonathan Livengood (2021). Causal Attributions and the Trolley Problem, *Philosophical Psychology*, 34 (8), 1167–1191.

- Sytsma, Justin, Jonathan Livengood, and David Rose (2012). Two Types of Typicality: Rethinking the Role of Statistical Typicality in Ordinary Causal Attributions, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43, 814–820.
- Sytsma, Justin and Edouard Machery (2010). Two conceptions of subjective experience, *Philosophical Studies*, 151 (2), 299–327.
- Sytsma, Justin and Kevin Reuter (2017). Experimental Philosophy of Pain, *Journal of Indian Council of Philosophical Research*, 34 (3), 611–628.
- Sytsma, Justin and Melissa Snater (forthcoming). Consciousness, Phenomenal Consciousness, and Free Will. In P. Henne and S. Murray (Eds.), *Advances in Experimental Philosophy of Action*, Bloomsbury.
- Thagard, Paul and Richard Nisbett (1983). Rationality and charity, *Philosophy of Science*, 50, 250–267.
- Turney, Peter and Patrick Pantel (2010). From frequency to meaning: Vector space models of semantics, *Journal of Artificial Intelligence Research*, 37, 141–188.
- Ulatowski, Joseph, Dan Weijers, and Justin Sytsma (2020). Cognitive Science of Philosophy Symposium: Corpus Analysis, The Brains Blog
<https://philosophyofbrains.com/2020/12/15/cognitive-science-of-philosophy-symposium-corpus-analysis.aspx>
- Urmson, J.O. (1969). A symposium on Austin's method. Part I. In K.T. Fann (ed.), *Symposium on J.L. Austin* (pp. 76–86). London: Routledge.
- Waismann, Friedrich (1997). *Principles of Linguistic Philosophy*. 2nd ed. Ed. by R. Harré. London: Macmillan.
- Webb, Taylor and Michael Graziano (2015). The attention schema theory: A mechanistic account of subjective awareness. *Frontiers in Psychology*, 6, doi:10.3389/fpsyg.2015.00500
- Weinberg, Jonathan (2015). Humans as instruments, on the inevitability of experimental philosophy. In E. Fischer and J. Collins (Eds.), *Experimental philosophy, Rationalism, and Naturalism* (pp. 171–187). London: Routledge.
- Weinberg, Jonathan, Shaun Nichols, and Stephen Stich (2001). Normativity and epistemic intuitions, *Philosophical Topics*, 29, 429–460.
- Weksler, Assaf, Hilla Jacobson, and Zohar Bronfman (2021). The transparency of experience and the neuroscience of attention, *Synthese*, 198, 4709–4730.
- Wittgenstein, Ludwig (1953). *Philosophical Investigations*. Oxford: Blackwell.
- Zahorec, Mike, Robert Bishop, Nat Hansen, John Schwenkler, and Justin Sytsma (forthcoming). Linguistic Corpora and Ordinary Language: On the Dispute between Ryle and Austin about the Use of 'Voluntary', 'Involuntary', 'Voluntarily', and 'Involuntarily'. In David Bordonaba Plou (Ed.), *Experimental Philosophy of Language: Perspectives, Methods and Prospects*, Springer.
- Zwaan, Rolf (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin and Review*, 23, 1028–1034.