A New Role for Experimental Work in Metaphysics

L. A. Paul

Abstract Recent work in philosophy could benefit from paying greater attention to empirical results from cognitive science involving judgments about the nature of our ordinary experience. This paper describes the way that experimental and theoretical results about the nature of ordinary judgments could—and should—inform certain sorts of enquiries in contemporary philosophy, using metaphysics as an exemplar, and hence defines a new way for experimental philosophy and cognitive science to contribute to traditional philosophical debates.

1 Introduction

Over the last thirty years, as the influence of logical positivism and neo-Kantianism have waned, many debates in contemporary philosophy have slowly shifted from analyzing our concepts about the world back to the more traditional task of using reason to determine facts about the world itself. There is now a great deal of active, new philosophical research that takes this world-focused perspective.

This return to the world is most evident in metaphysics, epistemology and ethics, but spans most of the philosophical subfields. The emphasis in these discussions has shifted from work on the explication of concepts, such as the analysis of our concepts of time, knowledge and goodness, to work on the explication of the nature of the world, such as the analysis of the nature of time, or an explication of what knowledge *is*. These are philosophical theories whose primary aim is to tell us what there is in the world, not to tell us what our concepts are of things in the world, nor to tell us what there is in the world via explicating our concepts of what there is in the world.

L. A. Paul (✉)
Department of Philosophy, University of North Carolina, Chapel Hill, NC, USA
e-mail: lapaul@unc.edu

The change in method is highly relevant for those who believe, as I do, that there is an important place for empirical work in philosophy. It is particularly relevant for experimental philosophers. While a role in contemporary philosophy for experimental philosophical data about the nature of our concepts has been actively developed and defended, the change in methodology from a focus on concepts to a focus on the world itself creates a need for new connections with cognitive science and experimental philosophy.

Although I'll develop this idea in more detail below, let me cut to the chase in order to get to the immediate point. The standard approach of experimental philosophy is to use empirical methods to test philosophical hypotheses about the nature of concepts. And with good reason: lots of philosophy in the twentieth century, especially in the mid-twentieth century, focused on evaluating our judgments about our concepts of things in the world, and empirical data here can be very helpful. (For a good example see Hitchcock and Knobe 2010.) But more recent philosophical work engages in a very different sort of project, one that reduces or eliminates the emphasis on concepts in favor of a more direct investigation about the nature of things, properties and other sorts of entities in the world. This more direct investigation replaces our intuitions or judgments about our concepts with intuitions or judgments about the natures of entities in the world, often drawn from our ordinary experience of them. In such work, a central place is preserved for quasi-perceptual judgments used in arguments from common sense and ordinary experience. This change in method entails that there is a need for philosophers to draw on cognitive science in a way that has little or nothing to do with concepts, and provides a corresponding opportunity for new experimental philosophical research. I'll develop this picture in more detail below. Please note: while this role for empirical results about ordinary judgments arises in nearly every sub-field of philosophy, I shall concentrate on one in particular, metaphysics, in order to keep my focus sharp.

2 Conceptual Analysis

In order to proceed, we must first get a sense of the different methodologies used by metaphysicians. Broadly speaking, there are two central methodological strands in contemporary metaphysical work: the first involves conceptual analysis, and the second, newer strand is based upon ontological explication and ontological reduction. The first strand descends directly from the twentieth-century focus on the analysis of metaphysical concepts and their associated terms. Metaphysics based on conceptual analysis reigned in the twentieth century, especially in the latter half of the period. The reasons for this are diverse, but are mostly grounded in the positivist rejection of traditional rationalist and other constructionist metaphysical programs. This rejection found expression in a range of stances, including the view that *a priori* analysis must be applied to our concepts rather than to the world, and the view that the proper task for philosophy is to determine what our conceptual framework commits us to believing about the world, not what the human-independent world actually contains.

In the conceptual analysis of metaphysical concepts, the focus is on analyzing concepts such as our concept of time or of causation, or of the meaning of the word "cause," and the hope, at least some of the time, is that an explication of these concepts or meanings will tell us what the world would have to be like in order for us to use the concepts and terms correctly. When doing conceptual analysis, one might only be interested in our ordinary use of such concepts and terms, or our ordinary commonsensical understanding of the entities the concepts and terms are about, in order to explicate our ordinary concept. Or one might be interested in constructing a philosophical concept that in some way derives from our ordinary concept yet is supposed to be an improvement upon it. One might think that *a priori* analysis simply acts as a refinement of ordinary concepts into philosophical ones, or one might wish to actually replace the ordinary concept with the philosophical one. In both approaches, attention is paid to our everyday or commonsensical understanding of the terms and concepts. Depending on the nature of the concept being analyzed, there is a recognizable (although its impact is hotly disputed) role for experimental psychological results. The role for experimental work is to determine the features and contents of the concepts, such as the concept of causation, of experimental subjects. The dispute focuses, in part, on whether the goal of conceptual analysis is to explicate the concepts of ordinary people or of trained philosophers, and whether *a priori* analysis of concepts accomplishes a different task from experimental determinations. Related issues come up with regard to connections between conceptual analysis and work in semantics and philosophy of language.

The most recent metaphysical work based on the analysis of concepts usually defines its project as one of determining what it is necessary for things in a world to be like in order for the metaphysical concept to apply to them. This is a project to determine what properties or constituents a world must include in order for a particular concept to refer to entities in that world. The idea is that we are trying to determine what must exist in order for the concept to have a referent in a world, which is a claim about what needs to be the case in any possible world in which the concept applies, not just the actual world. Often, the project is reductionist in the sense that it explicates one concept in terms of another, more fundamental concept, again in an analysis that extends across all possible worlds. A version of this approach is evoked by David Lewis when he says "I think conceptual analysis is required to reveal what it is that all the actual and possible varieties of causation have in common" (Lewis 2004, p.76).

One might then try to link up the results of the analysis with claims about what is actually in our world, i.e., one might try to link up one's conceptual analysis with an ontological theory of what there is. One might think that we know the concept refers in our world as the result of our direct acquaintance with the referent of the concept. (This seems to be Lewis's stance in his 2004.) Or one might argue that our best theories guide us to knowledge of the referent of the concept. For example, one might hold that when our philosophical and scientific concepts are successful (in some suitably defined sense) they pick out the things they describe in our world. If, after a successful conceptual analysis, we know what the applicability of the concept requires, and—because of the success of our theories involving these concepts—we know the concept actually applies,

then by extension we know something about what is in the world. For example, say we have successfully analyzed our concept of causation as a concept of the transmission of the conserved physical quantity of force. Now, if our successful physical theories involve our concept of causation, and we can infer truth from the success of these theories, we know that instances of causation are transmissions of forces and that there are such transmissions in our world. Another, less direct way to do this is to analyze the role the concept plays, following Lewis's (1970) treatment of theoretical terms, and then take the referent of the concept to be whatever most natural entity there is that comes closest to filling this conceptual role. Fans of the Canberra Plan (such as Menzies 1996) take this approach.

In any of these projects, the analysis of concepts takes a central role, and empirical work on the nature of the concept being analyzed is surely relevant. But these projects do not encompass all or even the most central metaphysical projects pursued by contemporary philosophers today.

3 Ontology

A more central metaphysical project is defined by the second methodological strand in metaphysics: ontology, which repudiates conceptual analysis as the end product. The primary task of an ontological theory is (usually) to develop an account of the nature and existence of the objects, properties, relations, events and other entities in the world. For example, an ontological theory of causation is a theory of what the causal relation is, or of what it might reduce to. Such a theory might hold that causal relations are (or in other words, reduce to) relations of counterfactual dependence, or are transfers of conserved physical quantities like force or momentum, or are instances of a prescribed sort of nomic subsumption. One's ontology need not be reductive: one could have an ontological theory of causation that took causal relations to be irreducible relations that are somehow characterizable as fundamental constituents of the world.

Ontology, thus described, is part of metaphysics, but is not centered on the analysis of concepts. On the ontological approach, conceptual analysis may be useful in order to help us refine our inquiry or direct our attention in the appropriate way, but it is at most a bit player in the real investigation. As I characterize it, then, an ontological theory tries to tell us about what the actual world includes, and perhaps about what other possible worlds include. Such approaches to ontology try to tell us about the world directly, rather than by routing the inquiry through conceptual analysis.

For those unfamiliar with this approach to metaphysics, it may be helpful to digress for a moment, as follows, to explain a bit more about the contemporary ontological stance. While ontologists have returned to the traditional metaphysical task of trying to determine the nature of the world itself, and in addition have returned to the use of *a priori* reasoning in order to do so, they have not returned to the traditional methods of their forebears, the old-style rationalists. For example, most ontologists do not rely on *a priori* theses about divine epistemology, innate ideas and infallible introspective reasoning in order to develop and justify their views. Instead, ontology has responded to the success of contemporary science by

emulating methods employed in the sciences when reasoning to conclusions. In other words, the contemporary ontologist makes use of *a priori*-based reasoning coupled with quasi-perceptual judgments ("intuitions") drawn from ordinary experience about the world to propose a theory of the way the world is, which we can understand as a proposal that a model or a class of models represents the world, and then relies on inference to the best explanation to draw conclusions.

This is the ontological stance taken by many contemporary (realist) metaphysicians. Just as with natural and social scientific approaches to constructing theories about the world, ontologists develop empirically adequate models of the world based as much as possible on our perceptual judgments and experiences and then evaluate them using theoretical desiderata such as simplicity, elegance, parsimony and overall theoretical fit. Central differences between ontology and research in the natural and social sciences include the emphasis on ordinary experience rather than, say, lab- or field-based studies, and the topics investigated. The difference in topics means that the empirical resources for ontologists are significantly more meager than those of their natural and social science counterparts, since much of what ontologists reason about counts as unobservable. (But note that, when appropriate, empirical work from physics or other sciences is, or at least should be, incorporated as a constraint on the model.) Although the empirical resources are more minimal than their natural and social scientific counterparts, the implicit idea is that just as in natural and social science, as long as we construct and evaluate our theories appropriately we are justified in inferring conclusions about the world based on empirical results culled from what we perceive. The comparative lack of empirical resources means that we have many more *prima facie* attractive candidates to choose from when evaluating empirically adequate models for their theoretical attractiveness, but this does not impugn the legitimacy of the model-building method, or of using theoretical desiderata like simplicity to evaluate a theory. I've developed further particulars of how we should understand ontological modeling and ontological inference to the best explanation in Paul (forthcoming).

So ontological theories can be understood as proposing metaphysical models of the world that represent it as including particular sorts of entities or as being a certain way. Claims such models make include claims about reductions of one sort of entity to another, claims about eliminability, and inflationary claims about what exists. For example, some ontologists argue for the thesis that objects are reducible to collections of property instances. Some argue for eliminations, for example, that there is no relation of composition, and hence there are no composites in the world. Some argue for inflations, for example, that the world contains emergent mental properties or irreducible *de re* modal features. Sider (2009) discusses the development of an ontological theory as a procedure where one develops an account of what the quantificational structure of reality is, and takes realist ontological disputes to be disputes about what the most natural satisfiers are in the world that fit that structure. The account of the ontological method I am describing can include work done under this more abstract rubric: on my view, Sider is taking a certain sort of work in ontology to involve the construction of models of the quantificational structure of the world.

So let's return now to the main point. The picture of ontological modeling I've been developing thus far has two main components. First, the models are proposed

as empirically adequate representations of parts of the world, and second, the models are evaluated using theoretical desiderata to determine the best candidate theory. Those who are skeptical of the value of doing metaphysics will probably want to dispute the legitimacy of the inference involved in the second component. They may also dispute the assumption involved in the first component, that ordinary experience gives us any knowledge at all about the entities ontologists wish to investigate. But since this discussion is not concerned to defend the overall value of doing ontology, I want to leave these disputes aside. The arguments I'm developing here are concerned with the implications for those who accept the first component, in particular, with the implications of relying on empirical evidence from ordinary experience to construct models.

The relevant idea I want to examine is not whether it is acceptable to use ordinary experience to construct models, but which features of ordinary experience should be used when constructing these models. Let me first say a little more about how ordinary experience plays a role in metaphysical modeling, and then, in a subsequent section, I'll look more closely at the question of what sorts of experience should be used.

Some ontological theories, such as those about the nature of spacetime, the direction of processes at the microphysical level, dispositions, or scientific laws, draw on well-accepted theories in physics and other natural sciences to construct models. But often, ontological theories need to go beyond what natural science has given them, as they engage with parts of the world that are ignored, assumed or otherwise not discussed explicitly by natural scientific theories. Such theories include those concerning part-whole relations, *de re* modal facts about identity, necessity and contingency, or theories about ontological fundamentality or priority. The source of information for the models here is different from that we can glean from the natural sciences, since it is based on ordinary experience. Often, when ontologists draw on ordinary experience to construct a model of something in the world, the ontologist starts out with the assumption that, by having experience of the world (and the things in it) we have some pretheoretical knowledge of this thing. The thought is that our everyday observations and experience give us this knowledge: it is supposed to be recognized by any suitably alert and thoughtful observer. To the extent that we have such knowledge, we have empirical data on the subject of enquiry. This is why it is a mark against an ontological theory if it is inconsistent with our common sense judgments or ordinary experience. At the very least, the theory must explain why such inconsistencies are not evidence against the theory. As a result, while the sources for ontological models include theoretical and empirical results from natural science, especially physics, in many cases ordinary knowledge is the central source.

Now, so as not to prejudge the case, since knowledge is a success term, let's describe propositions stating claims about pretheoretical knowledge about the world that are based somehow on ordinary experience as our "ordinary judgments" about the world. Often, ontologists take ordinary judgments to be constraints on their models. Above, I described these ordinary judgments as "quasi-perceptual," meaning to emphasize that, first, the content of these judgments is open to a certain amount of interpretation, and second, while some of the judgments assume that we are detecting empirical facts about the world using our sensory apparatus, they are

not confined to being about such facts. Indeed, many of the judgments seem to rely on the idea that we are able to make inferences concerning the nature of the world that extend past what we detect with our senses, although such inferences may be based, at least initially, on sensory detections.

For example, Hawthorne and Sider (2002), in a paper on the bundle theory of objects, say "We will examine modal considerations on reifying locations. Denying the existence of spatiotemporal locations excludes certain possibilities for spatio-temporal reality. Denying the existence of qualitative locations excludes certain possibilities for qualitative space. In each case the excluded possibilities are pre-analytically possible. Some of the possibilities can be reinstated by modifying the locationless theories, but at the cost of an unattractive holism. Do these modal considerations mandate postulating locations? That depends on whether modal intuition can teach us about the actual world" (p. 53). And Markosian (2008), discussing mereological composition, holds that "I think it is fair to say that according to common sense, it is not the case that for any group of objects, there is automatically an additional object composed of the members of that group. That is, according to common sense, composition is restricted" (p. 341). Paul (2006) sets up a discussion about coincidence and constitution with the claim that "[t]he intuitively appealing position is pluralism: it is natural to think that the statue has a context independent nature or essence that is different from the piece of clay, and thus that the statue is a different sort of thing from the clay. It is natural to think that as a result of deep, objective differences in the sorts of things they are, the statue can be destroyed while the piece of clay is preserved, e.g., when the piece of clay is molded into a brick" (p. 624).

The considerations and "pre-analytical possibilities" described by Hawthorne and Sider seem to be determined by modal intuitions derived somehow from ordinary judgments about the actual world. Markosian is more forthright about his reliance on ordinary judgment: he is concerned to describe a claim about when mereological composition fails to occur, and explicitly draws this claim from our ordinary judgments about what objects do and do not exist in the actual world. (Composition is not perceived in any obvious way: a plural collection of parts arranged in a certain way will not look any different from the whole that the parts compose.) Paul's claims about what is "natural" to think are simply claims about our ordinary judgments about essences and natures. All of these judgments are based on but extend past what we would normally class as perceptual experience.

I think such claims are perfectly legitimate as starting points for ontological modeling. After all, *any* theory of the world must start with our reflections and considerations about the way the world is, and every theory is infected with judgments based on our experience, even if many such judgments are ultimately dispensed with. This is true even for theories in natural science—including those of fundamental physics—for when such theories connect with the world through empirical work, we use results gleaned from what we perceptually detect (even when using instrumentation) to construct and evaluate our scientific models.

So metaphysical modeling of the world makes room for ordinary judgments about the world, and I maintain that, in principle at least, the use of such judgments is legitimate. We can characterize the use of ordinary judgments in ontology more generally. Arguments from ordinary judgments are used as premises for theories, motivations for a particular approach to a problem, reasons for the interest of an

ontological puzzle or question, or constraints on a successful theory. Arguments from ordinary judgments are also used to reject or cast doubt on competing theories. Sometimes the appeal to ordinary judgments is made explicitly, but more often it is implicit and couched in the form of a claim about intuitions, which are sometimes developed or marshalled using a thought experiment. Of course, ordinary judgments are regarded as defeasible. An ontological model that includes arguments from ordinary judgments usually develops an account of what objects the world contains, and then interprets our scientific and ordinary claims about the world in light of its representation of what there is.

Once we have a range of ontological models of how the world might be, as I described above, we choose between them using *a priori* reasoning based on theoretical desiderata such as overall simplicity, explanatory power, fertility, elegance, etc. In most cases, an acceptable theory must maximize theoretical virtues while being empirically adequate. What does empirical adequacy amount to here? It amounts to presenting a theory of what exists in the world that can make sense—even in a revisionary sense—of what we take to be our ordinary and scientific understanding about the world. Thus, as I've described the ontological method here, there is a significant and well-understood role for claims drawn from ordinary knowledge in many ontological theories. An ontological model that represents the world in a way that is inconsistent with our ordinary judgments, especially one that is radically inconsistent, is usually taken to have an especially heavy explanatory burden it needs to discharge, perhaps by maximizing its other virtues.

Let's recap, just to make the point absolutely clear, why ordinary judgments are included in contemporary ontological theories. First, the goal of work in ontology is to develop a theory about the nature of the world, and the method involves the construction of models that defend a certain account of what exists in the world. In constructing these models, ontologists assume that a main route to knowledge about the world is via knowledge gained from experience. Some of the content of this experiential knowledge is captured by natural scientific theories, but some of it is captured by our ordinary judgments, which are sometimes cashed out in terms of wanting to preserve commonsensical theories of the world. So part of constructing an ontological model involves drawing on what we think we ordinarily know about the world, and part of evaluating the success of an ontological model involves evaluating how well its representation of what exists fits with our ordinary judgments about the world.

4 Example: Theories of Causation

I'll flesh this out by looking at an example. Consider a topic that has recently attracted a reasonable amount of attention: ontological treatments of macro-level causation. The ontologist who develops a theory about what causation is, is developing a model of the causal relation. The model might try to represent causation in the actual world, or it might be more ambitious and try to represent causation in any possible world.

Assuming, as most ontologists do, that we have causation in our world and that we are able to experience and detect some cases of token causation, to build a model

of causation in the world the ontologist starts with a list of our ordinary judgments about putative cases of causation between events related in certain ways.[1] By inspecting this list, she can use it as well as other ordinary judgments (along with evidence drawn from natural science) to determine some of the characteristic features of what we take to be instances of actual or possible causation. She may then generalize and argue, for example, that certain of these features are necessary for causation, i.e., that they capture the nature of the causal relation.

The resulting model will be evaluated in part on how well it captures our ordinary judgments about causation. For example, does it represent a ball shattering a window as the ball causing the shattering? Can it handle omission-involving causation? Does it represent my failing to drink any coffee as a cause of my crankiness? Does it represent causation as transitive? If I kick the ball to you, and you kick it into the goal zone, was my kick among the causes of the increase in our score? A model of causation that can't accommodate many of our ordinary judgments, for example, a model that tells us that omissions can't be causes or that causation is not transitive, needs to explain why our judgments seem to conflict with its representation of the world.

This role for ordinary judgments in our modeling of causation is important and relevant, given two major assumptions. First, one must assume that these ordinary judgments about the nature of causation have been appropriately generated. Second, one must assume that the ordinary judgments about causation serve as a repository of (at least some of) what we know about causation.

Ontological work on causation that proceeds in the way I've described above is hardly unusual, and similar approaches to a wide variety of ontological topics (the nature of time, mereological composition, persistence, *de re* modality, etc.) are very common. Across these and other topics, developing a model of the world that employs ordinary judgments is a standard trope. The generic versions of the assumptions I described above, first, that we assume ordinary judgments have been appropriately generated, and second, that we take them to be a repository of what we know about the world, are ubiquitous. I'll now take a close look at the first assumption; that ordinary judgments have been "appropriately generated," which will lead us to the central point of this paper: that we need to consult experimental work when determining which ordinary judgments should be used in ontological models.

5 A Role for Cognitive Science in Metaphysics

What does it mean to say that we assume our ordinary judgments have been appropriately generated? At a minimum, we assume that the ordinary judgments we use in ontological models are the *ceteris paribus* judgments that normal subjects, taken to be reasonable approximations of rational agents with normal sensory capacities, would form under normal conditions. But, of course, this is not enough, for we must

---

[1] Depending on the stage of theory development, the ontologist may also examine ordinary or "folk" concepts as well as more philosophical concepts. In some cases the point of conducting this preliminary analysis of our concepts is simply to ensure that we are clear about the topic of discussion.

also assume that these ordinary judgments are not the products of illusions, and that they are constant across populations. (For example, the Müller-Lyer illusion is had by normal American subjects under normal conditions, but may be different in populations of subjects that have different (nonwestern) histories.)

This seems pretty obvious. It is equally obvious that one role for cognitive science in ontology is to identify places where our ordinary judgments might not be appropriately generated, as with illusions, and that ontologists must take note where any such judgments might crosscut judgments about ontological topics. More positively, ontologists would also be well advised to take note when ordinary judgments are confirmed as psychologically veridical. Although there is little evidence of such note-taking in either case, I believe this is because ontologists make the background assumption that, in fact, there are few or no such crosscuts, not because they don't see the relevance of the point. It would be nice if more attention was paid to this issue.

But never mind, since there is another way that results from cognitive science should play a role in our evaluation of ordinary judgments: one that is subtler to detect and define, and one in which there are empirical results, and where these results do crosscut ontological topics. The sort of thing I mean to be characterizing is the way that various environmental facts about a particular situation can act as non-veridical *cues* or *camoflages* that create a perceptual or quasi-perceptual bias in the observer for or against a particular conclusion.[2] Philosophers need to recognize this possibility in general, and need to know about particular possible cues or camoflages when evaluating ordinary judgments, since our judgments might be skewed as a result.

The claim is not that philosophers cannot draw on ordinary judgments when constructing their ontological models; the claim is that we must not do so in a naïve way. If we actually take ourselves to be developing a model of the way the world is given our best assessment of ordinary experience, we need to take into account current psychological knowledge about our interpretation and judgments of perceptual experience, whether such experience is imagined as part of a thought experiment or whether it actually occurs.[3] We need to know if there are perceptual quirks or habits that we need to adjust for when we evaluate premises involving case studies, thought experiments, or other ordinary judgments. In other words, given the role of ordinary judgments in ontological modeling, we need to know if facts about our cognitive apparatus result in certain sorts of perceptual or judgmental bias. How do we know this? The best way is to know if there are any relevant psychological or experimental philosophical results concerning the types of judgment being used.

The role for experimental work I am advocating is not one that is intended to replace current approaches to metaphysics, but to refine and strengthen such

---

[2] Here I am using terms based on Choi and Scholl (2006).

[3] To its credit, much recent work on causal modeling engages explicitly with empirical research. But it isn't ontology as I've described it—it isn't even reductive—and so it isn't the sort of work I'm targeting. A quote from Alison Gopnik (2007) can help us to see the contrast between the approaches to causation. "… causal graphical models are to causation as geometry is to space. Rather than providing a reductive definition of causation they instead provide a formal mathematical framework that captures important regularities in causal facts, just as the mathematical structure of geometry captures important spatial regularities" (p. 3).

approaches in much the same way that connections with current empirical work in physics and other natural sciences refines and strengthens metaphysical theories of the world. To make this more concrete, let me return to my example of causation. Albert Michotte (1946) performed a series of famous experiments that showed how individuals have the perceptual impression of causation under a range of different visual conditions (and also how certain environmental features removes the impression of causation). Michotte's work focused on "launching events," but subsequent researchers have developed a host of studies exploring the way that properties such as continuity of motion and relative changes in objects affect our causal judgments.

It is worth describing a version of Michotte's approach for those who are not familiar with it. Imagine yourself watching a computer screen with a stationary green object on it. You then see a red object move across the screen until it is adjacent to the green object with no space in between (so they appear to "just touch"); the red object then stops and the green object starts moving in the direction the red object had been moving. Merely imagining this experience should be enough for you to have the experience that Michotte discovered in subjects: the experience of the impression of causation, an impression that the red object causes the green object to move.

Saxe and Carey (2006) tell us that "[b]ased on his experimental results, Michotte proposed the existence of a special mechanism in the mind that transforms privileged inputs—visual sequences of motions with certain spatiotemporal parameters—into a 'genuine causal impression.' This mechanism exhibited the hallmarks of a perceptual process, rather than a conceptually mediated inference. The perception of causality in these events is fast and automatic, and subjects do not have introspective access to intermediate stages of the computation" (p. 146). In other words, the causal impression is perceptual rather than inferential.

Why should any of this be relevant to the philosopher of causation? Well, the obvious relevance of the work is that, since we are interested in ordinary judgments about causation, ontologists should have a clear sense of the properties of the situations that matter when we make such ordinary judgments. After all, part of our method involves attempting to generalize from ordinary judgments to determine what we know about which properties need to be instantiated to have cases of causation. For example, can there be causation across spatial gaps? Can there be causation across temporal gaps? Is causation invariably transitive? Does it require counterfactual dependence of the later event on the earlier event? Knowing what psychology can tell us about the properties we use to judge that there is causation is surely relevant.

But there is a less obvious point buried here. Think back to the description of Michotte's launching experiment. Recall that subjects, when they view a launching event, report having a causal impression, that is, they have the impression that they are detecting an instance of causation. As I described the experiment, it involved watching a computer screen on which images of objects moved—and hence, it involved watching a series of images appear at different places on the screen, not instances of actual causation. Yet: we still have the impression that there is causation involved in the motion of the images. So what Michotte showed (among other things) is that we can have the visual impression of causation even when there is no

causation involved, and even when we know there is no causation involved. The import of this point is that, if the perceptual cues are correct, we will judge that causation seems to be occurring. But there is more. Michotte didn't just show that we get an impression of causation, i.e., that it seems to us that we are perceiving, as we watch the images, that the green object causes the red object to move, even when there is no causation. He also showed that we will get the impression of causation *even when we know that* there is no causation. After all, we know that we are simply watching images on a screen.

The point here is that that we seem to be making some sort of brute perception-based ordinary judgment, but the impression does not correspond to real-world causal facts. This matters. A point about ontological work on causation will show why. The ontologist who wants a model of causation is in the business of constructing models based on our ordinary judgments of causation. When such models are reductive, as they often are, a central part of such models involves making "black-box" judgments about particular situations. In other words, a thought experiment or situation with certain features is described, such that the causal facts are not stipulated, and so the description is couched in entirely noncausal terms. Philosophers are then expected to evaluate the situation and judge whether a causal relation obtains. This is, in a sense, a gold-standard kind of case. If cases with a certain feature F are always judged to involve causation, the ontologist usually takes this to be evidence for the conclusion that feature F (counterfactual dependence, spatiotemporal contiguity, etc.) is sufficient for causation.

But if, as the Michotte results suggest, there are certain cues F that trigger an impression of causation independently of whether there really is causation, such as a certain sort of spatiotemporal contiguity, and if the black-box cases involve F, then the ontologist does not have the evidence that she thought she had. Another way to put the point: we need to know what the cues and the camoflages are, else we run the risk of mistaking cues for indications of causation, and missing causation because of camoflages. This is even more of an issue when we consider the fact that many of the cases ontologists develop far outrun everyday causal judgments in their complexity and structure, and so could be particularly vulnerable to judgments infected by perceptual bias.[4]

A different sort of example such that certain facts about it are epistemically underdetermined shows us another way we must be careful with respect to Michottean cues. The example involves the central way that the causation literature characterizes causal overdetermination. Causal overdetermination is characterized as involving cases where we have two events, A and C, each sufficient by itself for the effect E, and the way in which E occurs differs not at all from the way in which it would have occurred, had but one of A or C caused it. In the causation literature, overdetermination is often depicted using a causal "neuron" diagram like that of Fig. 1.

Here, the diagram represents that the signals from C and A are transmitted to E simultaneously. Each signal is sufficient, all by itself, to cause E, and E occurs.

---

[4] Tania Lombrozo (under review) has done some interesting experimental work on the complex causal cases that philosophers describe as "double prevention."

Fig. 1 A causal neural diagram of causal overdetermination
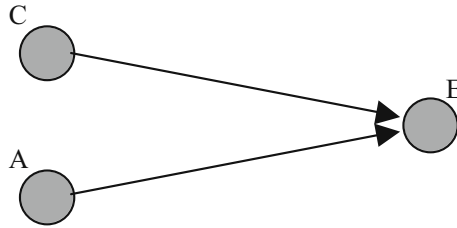
Such diagrams are enormously useful, and they are widely utilized in the philosophical literature on causation. Indeed, in the case of causal overdetermination, they may be essential, since an adequately precise definition of overdetermination does not exist. (For in-depth discussion of this problem see Hall and Paul, forthcoming.) But we must be careful when using such diagrams— or when using any other depiction of cases—that we are not unwittingly building Michottean cues into our story.

The problem with such cases of (putative) overdetermination is that whether both A and C are causes, and assuming they are causes, whether they are equal causes, is epistemically underdetermined. The neuron diagram above strongly suggests that both A and C are equal, full causes of E. But the characterization of over-determination I gave where "A and C are each sufficient by itself for the effect E, E occurs, and the way in which E occurs differs not at all from the way in which it would have occurred, had but one of A or C caused it" applies equally well to certain cases where only A is a cause, cases where only C is a cause, to cases where the combination A+C is a cause (but it is false that each of A and C is a cause; i.e., only their sum causes E, not each part acting separately), or even to cases where neither A nor C cause E.

Now, if we use Fig. 1 as a representation of the characterization I gave of (putative) overdetermination, we might be led by the fact that the arrows "just touch" the filled-in circle to think that, *obviously*, both A and C are full causes of E. The arrows touching the circle might be a Michottean cue. And hence, we might think we can "see" that A and C are full causes of E even if we lack a concept of how causation could possibly work this way (for example, prior to looking at Fig. 1, one might think that if A were to cause E, then it would necessarily exclude C from causing E, or vice-versa, since causation involves a sort of "complete making-happen"). Relatedly, one might think that the diagram shows that causal over-determination with both A and C as full, equal causes of E must be metaphysically possible, since we can *see* how it works using the diagram. The cues make it seem like we are detecting the possibility of causal overdetermination, but this could be the case even if causal overdetermination were metaphysically impossible. Another way to put the point is in terms of the representation: we assume Fig. 1 represents overdetermination. But if we were really just relying on Michotte-style cues when we had our causal impressions, there would be no justifiable reason to think that we had any special intuitive insight into whether Fig. 1 represents a case of overdetermination or, say, a case of preemption where A causes E and preempts C. (Hitchcock 2003 argues that neuron diagrams are often misleading in the way they represent causation.)

Another example comes from Scholl and Nakayama (2002) and Choi and Scholl (2006). By modifying Michotte-style cases, Scholl and Nakayama (2002) describe other ways perceptual bias can affect subjects' causal impressions. What they show is that goings-on that are extrinsic to the events involved in the putative case of causation under examination can affect our causal impressions. In Fig. 2(a) and 2(b), we have a series of static frames, where the arrows represent the direction of motion.

In 2(a) we have a Michotte-style launching case, where in the third frame the moving black disk "collides" with the stationary white disk, and then in subsequent frames the white disk moves to the right while the black one remains stationary. In 2(b) the sequence of events is exactly the same, except that in the third frame the black disk fully occludes the white disk instead of appearing to collide with it. When presented with a sequence as in case 2(b), subjects describe the sequence of events as a "noncausal passing" of the black disk: the black disk is seen as passing over the white disk, swapping colors with it when there is full occlusion, and then continuing to the right as a white disk (while the disk that was formerly white is now black and remains stationary).

But when extrinsic states of affairs are built into case 2(b), subjects change their causal impressions. In particular, when a temporally synchronized canonical launching event such as the one represented by 2(a) is included along with the case represented by 2(b), subjects think that both cases are cases of causation, where the black disk somehow collides with a stationary white disk and the white disk then moves to the right, leaving the black one stationary. Figure 3 represents the case where extrinsic facts change subjects' causal impressions.

Imagine someone trying to evaluate whether causation is or is not an intrinsic relation. If they evaluated cases of real causation that had structure relevantly similar to those of Scholl and Nakayuma's three cases, the cases could be thought to provide support to the view that causation was not an intrinsic relation, since one's causal impressions can change if extrinsic facts are added. But as the experimental work from Michotte-sryle cases shows, we have these changes in causal impressions independently of whatever the truth is about the causal facts of real cases. The example also shows how we might misjudge a single case of causation if it occurs in certain contexts.

Results like these make it obvious that we need to be clear on different cues and camouflages that might affect our causal judgments. Note also that the misleading information does not have to take the form of cues. Take the ontologist with her black-box cases. It might be that the presence of certain F camouflage impressions of causation when those impressions would have been veridical. Then we could fail to
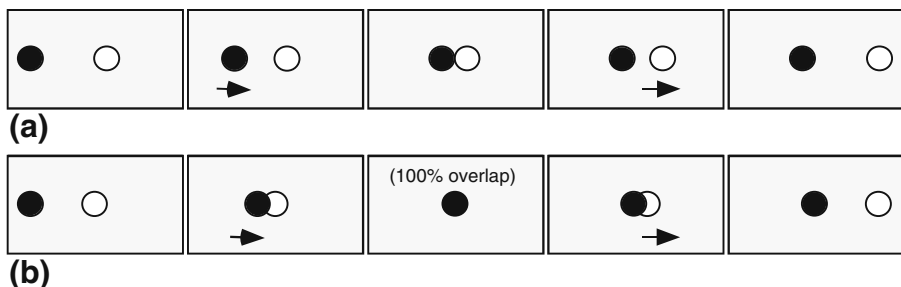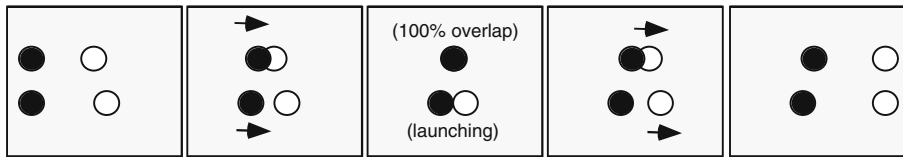
Fig. 2

Fig. 3

detect evidence that some fact was sufficient for causation, or was at least a property of causation, because our cases contained unnoticed causal camoflages.

Although I've focused on examples from work on the ontology of causation, my point generalizes. For example, psychological work on our perception of temporal sequences of events such as the phi-phenomenon and the flash-lag phenomenon is relevant to our ordinary judgments of the nature of time and persistence. See Le Poidevin (2007) for discussion of the relevance of the flash-lag phenomenon and Paul (2010) for an argument that cognitive results based on phi-phenomena are relevant to ontological accounts of temporal passage. Psychological work on the features we use to identify objects as persisting is surely relevant to ontological theories of material constitution and *de re* persistence properties, as well as to theories of what objects are, all of which draw motivation and insights from our ordinary judgments. See, for example, Xu (1997) and Xu and Carey (1996).

These are just some of the ways empirical work can be relevant to metaphysics. What I've been arguing here, in effect, is two-fold. First, there was a general methodological point about theorizing. I argued that philosophy is increasingly concerned with making and defending claims about the world rather than claims about concepts, and that philosophical theories in this vein can be understood as classes of models that purport to describe and represent the world. Second, there was a more specific point about the role of experimental work with respect to this sort of philosophical theorizing. Here, I argued that ordinary judgments are used to build and evaluate the models of philosophical theories of the world, and such judgments must be appropriately generated if we can use the models to infer conclusions about the world. The need for a determination of whether and how ordinary judgments are appropriately generated provides a new role for empirical work in contemporary philosophy.

Finally, two obvious points. Although I focused on ontology as opposed to conceptual analysis, to the extent that the conceptual analyst relies on ordinary judgments, she too must also evaluate ordinary judgments in light of the relevant psychological work. Also, as I indicated at the start of this paper, while metaphysics has been my exemplar, the use of ordinary judgments to construct models of the world is widespread in philosophy; it is used in epistemology and ethics as much as in metaphysics. Philosophy, broadly speaking, involves the construction of models and takes ordinary judgments to be constraints on such models, and hence needs to attend to the cognitive science of ordinary judgments.

References

Choi, H., and B.J. Scholl. 2006. Perceiving causality after the fact: postdiction in the temporal dynamics of causal perception. *Perception* 35: 385–99.

Gopnik, A. 2007. *and Schulz, L.* Oxford: Causal Learning.

Hall, N., and Paul, L. forthcoming. *Causation: A User's Guide*. Oxford: Oxford University Press.

Hawthorne, J., and T. Sider. 2002. Locations. *Philosophical Topics* 30: 53–76.

Hitchcock, C. 2003. What's wrong with neuron diagrams? In *Topics in contemporary philosophy vol. 4: Causation and explanation*, ed. J.K. Campbell and M. O'Rourke, 69–92. Cambridge: MIT.

Hitchcock, C. & Knobe, J. 2010. Norms and causation. *Journal of Philosophy*.

Le Poidevin, R. 2007. *The images of time: An essay on temporal representation*. Oxford: Oxford University Press.

Lewis, D. 1970. How to define theoretical terms. *Journal of Philosophy* 47: 427–46.

Lewis, D. 2004. Causation as influence. In *Causation and counterfactuals*, ed. J. Collins, N. Hall, and L. A. Paul. Cambridge: MIT.

Lombrozo, T. under review. Causal-explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions.

Markosian, N. 2008. Restricted Composition. In *Contemporary Debates in Metaphysics*, (eds.), Sider Hawthorne, and Zimmerman (eds.), Basil Blackwell, 341–64.

Menzies, P. 1996. Probabilistic causation and the pre-emption problem. *Mind* 105: 85–117.

Michotte, A. 1946/tr. 1963. *La Perception de la Causalité*, Institut Supérieur de Philosophie. English translation of updated. In *The Perception of Causality*, eds. T. Miles & E. Miles (ed.), Basic Books, 1963.

Paul, L. 2006. Coincidence as overlap. *Nous* 40: 623–59.

Paul, L. forthcoming. The Handmaiden's Tale: Ontological Methodology. In *Metaphysics and methodology*, ed. Ladyman James. Oxford: Oxford University Press.

Paul, L. 2010. Temporal experience. *Journal of Philosophy*.

Saxe, R., and Carey, S. 2006. The perception of causality in infancy. *Acta Psychologica* 123: 144–65.

Scholl, B.J., and K. Nakayama. 2002. Causal capture: contextual effects on the perception of collision events. *Psychological Science* 13: 493–98.

Sider, T. 2009. Ontological Realism. In *Metametaphysics*, ed. Manley Chalmers and Wasserman, 384–423. Oxford: OUP.

Xu, F. 1997. From lot's wife to a pillar of salt: evidence that *Physical Object* is a sortal concept. *Mind and Language* 12: 365–92.

Xu, F., and S. Carey. 1996. 'Infants' metaphysics: the case of numerical identity. *Cognitive Psychology* 30: 111–53.