The Method(s) of Cases

Jeffrey Maynes

Experimental philosophy has focused attention on the role intuitive responses to philosophical cases play in philosophical argumentation. The method of appealing to such cases has been dubbed the 'Method of Cases,' and in recent work, Edouard Machery has both defended its prevalence and uniformity in philosophical practice, and criticized its epistemic value. In this paper, I argue that there is no single Method of Cases, but rather a set of Methods of Cases. To defend this claim, I distinguish and articulate these different methods, and argue that they better explain several paradigmatic appeals to cases. This result not only challenges the homogeneity of the Method of Cases, but it stocks our methodological toolbox with additional interpretive tools which help us to not only better understand philosophical arguments, but to better understand the significance of experimental work.

1. The Method of Cases

A central point of dispute in recent work in philosophical methodology, driven by debates about experimental philosophy (X-Phi), is whether, and to what extent, philosophers rely on intuitions as evidence for and against philosophical theses. Arguments employing thought experiments are often classified as 'appeals to intuition' on the ground that they rely upon our intuitive response to the case described in the thought experiment. For example, in objecting to utilitarianism, I might ask you to imagine a 'utility monster' (Nozick, 1974), a person who gets enormous amounts of pleasure out of the pain of others. This person gets so much pleasure, in fact, that we would maximize pleasure in the world by catering to his or her whims, inflicting great pain on others. Obviously, or intuitively one might say, we should not do so, and since utilitarianism says the contrary, so much the worse for utilitarianism.

These cases, it is claimed, are rife in philosophy; paradigmatic cases include Socrates' counter-example to Cephalus' definition of justice in the Republic, Thomson's violinist case,

Kripke's Gödel-Schmidt case, and Searle's Chinese Room. In each, experimental philosophers argue, the crucial move is an empirical claim about what we (for some relevant set of individuals) find intuitive, for this intuition is used as evidence for or against the philosophical claim under consideration.

Mallon, Machery, Nichols & Stich (2009) dub the method of appealing to our intuitions as the `method of cases'. "It is difficult to overestimate the importance of the method of cases in contemporary philosophy," write O'Neill & Machery, "it plays a very large role in areas such as epistemology, metaphysics, ethics, philosophy of language, philosophy of mind, and action theory" (O'Neill & Machery, 2014, xiv). This method, it seems, is wide in reach, but also uniform in application. At least when cases are used evidentially, Machery argues for methodological homogeneity, or the claim that cases play the same role in areas ranging from epistemology to action theory (Machery, 2017). In this paper, I argue that the Method of Cases ought to be replaced with the Methods of Cases; that is, that cases are used in a number of ways in philosophy, and understanding these roles helps us to better understand debates over the significance of cases in philosophy, and the significance of results in experimental philosophy.

Definitions for the Method of Cases (MoC) vary from stringent to relaxed. A stringent definition is offered by Mallon, et al. in the context of the study of reference (Mallon et al., 2009, p. 338); put into a generalized form, this version of the MoC states that:

Method of Cases (Strong): The correct theory T for topic X is the one best supported by the intuitions of competent judges of X.

This formulation is strong in two respects. The first is that it restricts the relevant evidence to our intuitions about the topic (and so, the role of cases is to elicit intuitions). The second is that it

treats intuitive consistency as a *sufficient* condition on the correct theory. What it means to be "best supported" might be cashed out in a variety of ways, including the theory that accommodates the greatest percentage of our intuitions, captures the most important intuitions, captures the depth and complexity of our intuitions best, etc. Regardless, according to the Method of Cases, doing this better than the competing theories is sufficient to judge one's theory to be correct.

Such a strong formulation, however, is not our only option. In more recent work, Edouard Machery has worked with a more expansive conception of the MoC. In *Philosophy Within Its Proper Bounds* (Machery, 2017), he defines the method in terms of the *cases,* rather than in terms any particular mental state elicited by them. A philosophical case is simply a real or hypothetical situation put forward by philosophers. Put broadly, this version of the MoC is the use of philosophical cases to elicit some attitude (say, a judgment), which in turn is taken as evidence for the content of that attitude. This can be in the formal mode (using the case to draw inferences about the meaning of a term) or in the material mode (using the case to draw inferences about the referent of a term). For this paper, I will define the expansive formulation as follows:

Method of Cases (Expansive): Elicited attitudes A to a philosophical case C reliably indicate the truth of the contents of A, and those contents are used as evidence for or against a philosophical thesis T.

It is expansive in the sense that it is not restricted to any particular attitude (such as intuition), and that it does not contain any specific epistemic prescriptions about the relationship between the response and the correct theory. I will argue that the strong definition is implausibly strong to be of much use in the study of philosophical methodology, but also that the expansive

formulation is similarly too restrictive to account for the roles that cases play in philosophical methodology.

The shift from intuition-talk to neutrality about the elicited attitude may at first glance appear to be a substantial shift between the two definitions. The definition of 'intuition' (which have been defined as everything from intellectual seemings (Bealer, 1998) to mere judgments about the world (Devitt, 2006)) is typically accompanied by an account of how those intuitions provide evidence in philosophy and other disciplines, and what kinds of claims they can provide evidence for. However, the term "intuition" has also been used by experimental philosophers as a kind of catch-all for the responses elicited by thought experiments, *whatever* they may turn out to be. As such, we need not read the strong formulation of the MoC as being committed to any particular theory of intuition, only that the right theory is determined by consistency with the right set of intuitions.

The shift away from intuition talk is useful, and though in past work I have similarly used "intuition" to refer to the responses elicited by cases, doing so has arguably produced more confusion than benefit, as it suggests more homogeneity in the role of our responses cases than actually exists. Further, the philosophical significance of X-Phi, for both its critics (Cappelen, 2012) and its defenders (Alexander, 2012), has been tied to the significance of intuition. Yet, if used merely as a placeholder for 'whatever attitudes are elicited in response philosophical cases,' we can dispense with intuition talk altogether and instead ask whether any of the attitudes so elicited (whether they be judgments, intellectual seemings, or some other attitude) are used as evidence in philosophy, and if so, under what conditions. If so, then we can similarly disentangle the significance of X-Phi from the significance of intuition, and instead examine the role that cases and our responses to them play in various contexts. Further, experimental work typically

does not constrain respondents to only providing one particular kind of intuition or judgment (which may be fast or slow, reflective or non-reflective, rooted in certain kinds of phenomenology or not, etc.). Here, I will talk of *responses to cases* instead of intuitions (except in describing the views of others).

In order to defend the claim that there are multiple methods of cases, I will first offer a taxonomy of three types of appeal to cases: the reliabilist interpretation, the etiological interpretation, and the non-evidential interpretation. Within each of these types, I then identify several variations on the role cases can play. This establishes that there are multiple possible methods of cases. To make the further claim that philosophers in fact employ these methods, I offer supporting examples, which, I argue, are best interpreted according to these various methods. Each of these cases is philosophically rich, and would merit much more discussion than I have room to offer here. Nevertheless, it is my intention to show that the fit between these examples and the methods proposed is sufficient to show that we are better off rejecting methodological homogeneity, and recognizing a range of roles for cases to play. It is my hope that further work can make use of these interpretive strategies to further explore these, and more, examples in the philosophical literature.

Put in another way, it is my aim to add more tools to our methodological toolbox. We can stock our toolbox by describing a range of argumentative strategies that philosophers use, and we can use that toolbox to interpret arguments, and to differentiate between arguments that are superficially similar but nevertheless distinct in important respects. "I suppose it is tempting," Maslow writes, "if the only tool you have is a hammer, to treat everything as a nail" (Maslow, 1966, p. 15). If one's methodological toolbox is limited, then one is more likely to interpret arguments in conformity with those limited tools. The better we understand the diversity of

methodological approaches taken by philosophers, the better we can understand their arguments, making it less likely that we will shoehorn an argument into an inappropriate methodological description.

2. An Inventory

Machery identifies four roles that philosophical cases play. Two we have already seen:

Material MoC - Cases elicit responses about the referent of a relevant term. These responses are reliable indicators of the truth of their contents. Therefore, from responses to cases, we can infer facts about the referent of the relevant term.

Formal MoC - Cases elicit responses about the meaning of a relevant term. These responses are reliable indicators of the truth of their contents. Therefore, from responses to cases, we can infer facts about the meaning of the relevant term.

Two others are non-evidential:

Provocation - Cases play a provocative role when a conflict between the elicited response and existing theory, or between elicited responses to similar cases, motivates further inquiry.

Illustration - Cases play an illustrative role when the case exemplifies a philosophical claim, concept, or distinction, and is used for the pedagogical purpose of helping the audience better understand that claim, concept, or distinction.

I will treat the Formal and Material instances of MoC as sub-types of a broader category, which I will call the *Reliabilist* approach to the MoC. Setting aside the non-evidential roles, Machery argues that all instances of the MoC are instances of the reliabilist approach.. He

contends that "philosophers' material or formal use of cases is homogenous" (Machery, 2017, p. 29) on the grounds that philosophy is ultimate a dialectical activity, and if the use of cases did vary in its methodology, it would be unable to play this dialectical role (e.g., other philosophers would be in the dark about how to interpret one's arguments). This puts the burden of proof on the advocate for heterogeneity, and "in the absence of actual evidence that both the formal and material use of cases really differ across philosophers, research areas, or research traditions, we can ignore the remote possibility that the assumption of methodological homogeneity is false" (Machery, 2017, p. 29).

Construed narrowly, Machery's claim is simply that the material and formal methods are applied in the same way consistently. That is, there is no substantial variation in how users of the formal MoC apply the formal MoC. Construed more broadly, Machery's claim is that there is (for the most part) uniformity in philosophical practice, and that this uniformity is explained by the claim that uses of cases all (or almost all) fit one of the two reliabilist interpretations. It is this broader construal that I take issue with here. In the remainder of this section I will detail a series of alternative interpretations of the role that cases play in philosophical argument and discourse, and, in so doing, I will take up the burden of proof laid down by Machery.
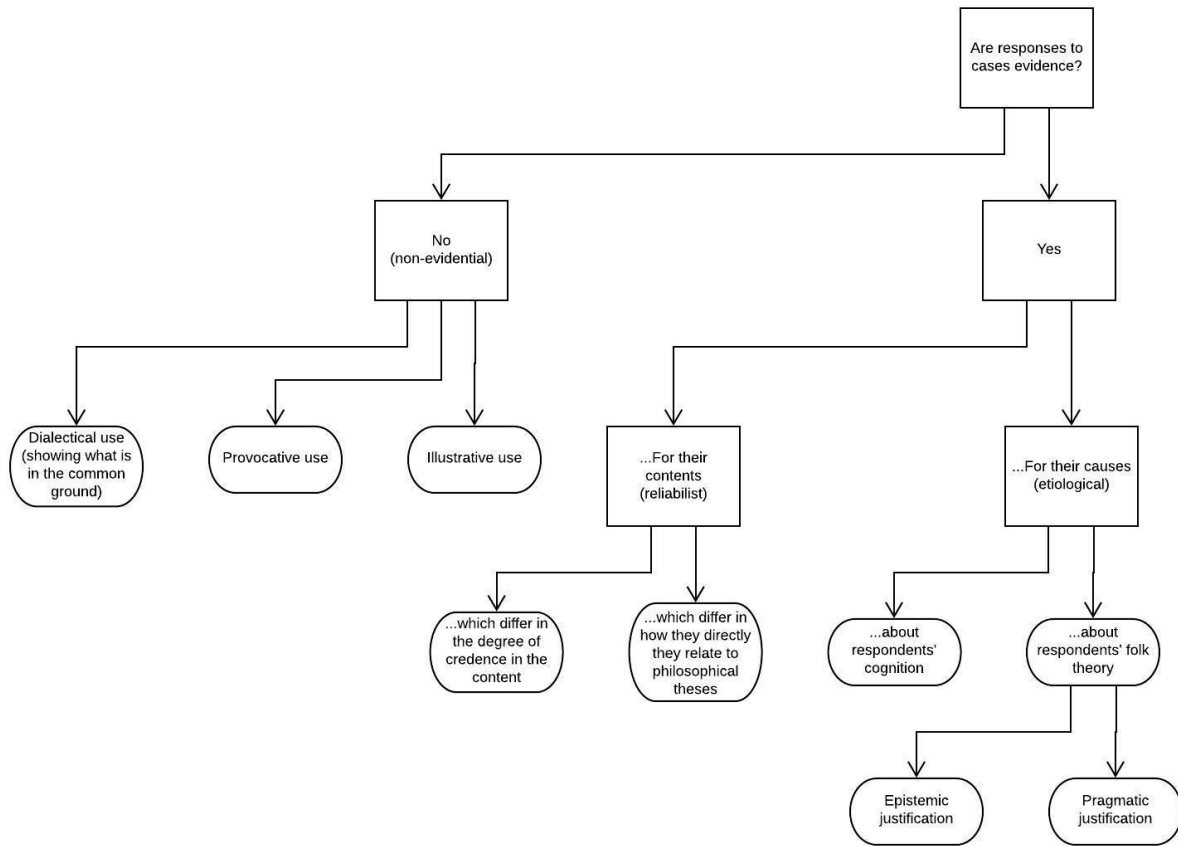
I distinguish three broad categories for the role of cases, each containing several sub-types. The three categories differ in how they answer the central question: is the response to the case used as evidence for the truth of some further claim? The first category, the reliabilist approach, answers that responses serve as evidence for the truth of their contents, based on the reliability of the respondent. To use Lycan's (1988) distinction with regard to intuitions, we infer the truth of the *intuited* content from the occurrence of the *intuiting* (the psychological act of having the intuition). I then distinguish sub-types within this category based on the degree of

credence the response lends to our belief in the truth of the content, and the way in which the content is then used in philosophical argumentation.

The second category, the etiological approach, answers that responses to cases are used as evidence, but not for the content of the response. Philosophers using a case in the etiological approach infer some claim about the factors leading to that response such as claims about the respondent's psychology, or prior theory. These claims may then be used in a variety of roles in philosophical argumentation. Unlike the reliabilist approach, the etiological approach need never presume that our responses actually indicate the truth of their contents. For example, that a respondent thinks it is wrong to push someone from a footbridge into the way of an out of control trolley car in order to save other lives may tell me about the psychology of the respondent, even if I draw no conclusions about whether the action itself is morally wrong.

Finally, the third category covers the the non-evidential use of cases, where the response itself is not used as evidence for or against any claim. Instead, the response plays some other role in helping us to understand a claim, recognize our prior commitments, or think about relevant philosophical questions. Put schematically, I develop the following taxonomy:

Are responses to cases evidence?

No (non-evidential)

Yes

Dialectical use (showing what is in the common ground)

Provocative use

Illustrative use

...For their contents (reliabilist)

...For their causes (etiological)

...which differ in the degree of credence in the content

...which differ in how they directly they relate to philosophical theses

...about respondents' cognition

...about respondents' folk theory

Epistemic justification

Pragmatic justification

2.1 The Reliabilist Approach

The form of inference in each of the reliabilist interpretations is:

1. Respondent(s) R has response p.

2. Respondent(s) R are reliable indicators, to degree of credence c, with regard to p.

3. Therefore, p, with credence c.

The advantage of such a straightforward schema is that it mirrors how we treat everyday judgments. I bring an umbrella with me on the walk to work when the meterologist predicts rain. I make this decision on the basis of her reliability, even without my having consulted (or being able to consult) the data which she used to make the prediction. In the Formal MoC, it is

presumed that we are reliable indicators about the meanings of terms, and in the Material MoC, that we are reliable indicators about the referents of our terms. That is, we are reliable judges about what we mean by "knowledge" or about the moral facts in trolley cases.

The grounds for such reliability in philosophy are contentious. The meteorologist has some expertise acquired through training and experience, but I am also able to empirically determine her reliability by comparing past predictions to actual weather. In philosophical cases, this latter option is closed to us; as Cummins (1998) notes, philosophical judgments cannot be calibrated, as we lack independent access to their subject matter. Perhaps philosophers are experts because we have special knowledge of philosophical concepts or theories (Ludwig, 2007), or general skills that enable us to parse and think through cases (Williamson, 2011). Perhaps certain kinds of responses themselves carry epistemic weight, such as a rational seeming of a necessary truth (Bealer, 1998). Or perhaps simply as competent users of a language, we all are reliable judges about the applicability of the common terms of that language.

This is a well-trod literature, and I will not focus on it here. Instead, I will focus on a less noticed distinction between reliabilist uses of cases, namely, the degree of credence assigned to our belief that *p.* At the strong end, one might interpret our reliability as establishing with certainty, or near enough, the truth of the contents of our responses. Such a strong interpretation would square with a counter-example interpretation of the MoC, where the case elicits a response that in turn strongly establishes a fact which serves as counter-example to the prevailing theory. For example, if the Gettier case establishes that the character in the case has justified true belief but not knowledge, then this shows that justified true belief *cannot* be knowledge.

Yet, the use of cases is often more restrained. Consider, for example, Ned Block's use of the China-Mind case to challenge functionalist theories of mind (Block, 2007a). A crucial feature

of functionalism, indeed, one of the famous arguments in favor of it (Putnam, 1975), is that

mental states are multiply realizable. Block argues that this feature of functionalism leads to fatal

difficulties. If we accept the view, we are forced to either attribute minds to things without a

mind (the liberalism objection), or we have to deny minds to things that have them (the

chauvinism objection). To show how functionalism falls prey to liberalism, Block asks us to

imagine a situation where the citizens of China instantiate an extensive Turing machine that is

functionally identical to a human mind. Each individual in the system carries out a specific task

(in the same way that a specific component of a computer might, or a neuron) and the results of

this system are transmitted to a human body. The body behaves in the same way a normally

minded individual would, including yelling "ouch!" when s/he stubs his/her toe.

Block asks: does the China-mind (constituted by the body and the Turing machine) feel

pain when it stubs its toe? Does it have qualia? The answer, he contends, is obviously that it does

not. Since such a system is functionally identical to a human mind, it must, if functionalism is

true, be a mind. Yet, it clearly is not identical to a human mind, for it lacks qualia.

One might be tempted to read this as a straightforward counter-example, where the case

establishes that a system functionally equivalent to a human mind would not have qualia. Yet,

Block cautions against such a simple reading:

> My case against functionalism depends on the following principle: if a doctrine
>
> has an absurd conclusion that there is no independent reason to believe, and if
>
> there is no way of explaining away the absurdity or showing it to be misleading or
>
> irrelevant, and if there is no good reason to believe the doctrine that leads to the
>
> absurdity in the first place, then don't accept the doctrine. I claim that there is no
>
> independent reason to believe in the mentality of the [China-mind], and I know of

no way of explaining away the absurdity of the conclusion that it has mentality

(though of course, my argument is vulnerable to the introduction of such an

explanation). (Block, 2007a, p. 77).

In a critique of Searle's (seemingly similar) Chinese Room argument, Block notes that "the real

crux of the debate rests on a matter that Searle does not so much as mention: what the *evidence* is

for the formal-symbol manipulation point of view" (Block, 2007b, p. 104, emphasis in original).

Searle's mistake, according to Block, is treating our intuitions as being sufficiently strong

evidence that they could serve as a falsifying counter-example against Strong AI, regardless of

any evidence in favor of it. Block's use of the China-mind case does function similarly to the

other reliabilist cases. From the fact that we judge the China-mind not to have qualia, we are

licensed to infer that it, in fact, does not have qualia (provided that we cannot explain away the

judgment). Yet, we have a lower degree of credence in the truth of the content than in the

counter-example cases, which is defeasible in light of countervailing evidence.

Our responses to cases may provide even weaker evidence in favor of the truth of their

contents. In an experimental study on folk intuitions about consciousness, Nahmias et al.

describe their results as locating where the burden of proof lies:

> [I]f a philosophical theory does turn out to be privileged by the endorsement of
>
> the folk, that would seem to position the burden of proof on the shoulders of those
>
> who argue contrary to folk intuitions. If it turns out that a significant majority of
>
> people make judgments that support either compatibilist or incompatibilist views,
>
> that would at least give "squatters' rights" to whichever position has such support.
>
> (Nahmias, Morris, Nadelhoffer, & Turner, 2005, p. 564)

This notion of 'squatter's rights,' or of locating the burden of proof, can be understood epistemically or dialectically. In the latter sense, the folk intuitions do not offer reason to believe the position, but rather simply *describe* that it is the position currently believed or assumed to be true. This use falls under the common grounds interpretation discussed below. Understood epistemically, our responses to cases about free will provide evidence in favor of compatibilism or incompatibilism antecedent to uncovering further evidence. I will call this the *prima facie* instance of the MoC.

Thus far, I have considered reliabilist interpretations that constrain philosophical argumentation directly. That is, that the contents inferred from our responses bear directly on whether philosophical claims are true or false. An indirect approach is suggested by those who use cases to fix the object of inquiry. Michael Devitt (Devitt, 2006) and Hilary Kornblith (Kornblith, 1998) defend a role for intuition in helping us identify a target for investigation, but contend that intuitions play no subsequent role once better empirical methods are brought to bear. Let us call this the *weak topic-fixing* account.

Alvin Goldman defends a *strong topic-fixing* account, arguing that our intuitions play not only a topic-fixing role at the outset of inquiry, but that they maintain that connection throughout the study (Goldman, 2007). For example, "examining folk epistemic concepts should reveal how truth (true belief) is a primary basis of epistemic evaluation… when moving from folk epistemology to scientific epistemology, we should retain the concern with truth-related properties" (Goldman, 2007, 22). Philosophers using this method often use cases to explore concepts like knowledge by testing the conditions in which we are willing to apply that concept. More precisely, the strong topic-fixing account is both a linguistic claim, and an epistemological claim. The former is that the term for the object of inquiry (say, 'belief') operates as a

description composed of a cluster of central intuitive properties (which may be definite or indefinite, depending on whether uniqueness is one of those properties). If inquiry settles on a referent that does not possess those properties, then, even if it is a worthy of object of inquiry, it is not an instance of the original term.

The epistemological claim is that we should care about keeping inquiry focused on the original object of inquiry; that is, that such an inquiry is valuable enough that it should, at the least, be pursued in parallel with revisionary lines of inquiry that focus on objects lacking the intuitive properties. After all, one could accept the linguistic claim, but simply accept that science moves on from the vocabulary of common sense. An eliminative materialist, for example, is quite happy to abandon belief-talk on the grounds that what is actually going on in the brain is quite unlike our folk psychology. Yet, according to the strong topic-fixing account, it is understanding belief that was originally motivating inquiry, and even if understanding non-belief brain states is interesting and important, we also have an interest in understanding what beliefs are. Beliefs may not turn out to be exactly what we expect, but only if some of the intuitive properties are shared can we call it a theory of belief at all.

Consider David Chalmers' distinction between the two problems of consciousness (Chalmers, 1997). Of the easy problems, Chalmers writes:

> The word 'consciousness' is used in many different ways. It is sometimes used for the ability to discriminate stimuli, or to report information, or to monitor internal states, or to control behavior. We can think of these phenomena as posing the "easy problems" of consciousness. These are important phenomena, and there is much that is not understood about them, but the problems of explaining them have the character of puzzles rather than mysteries. (Chalmers, 2003, p. 103)

These problems stand in contrast with the hard problem, which is explaining how experience in all of its qualitatively rich glory arises from physical processes. That is, how do we explain the sweetness of eating a peach? It seems that one could explain how our minds identify the peach as sweet without thereby committing ourselves to our experiencing sweetness. In distinguishing these problems, responses to cases like eating a peach, or any of the host of cases used in the study of qualia, are used as topic-fixers. The responses in question are simply our judgments about the qualitative properties of experience, e.g., that eating a peach tastes sweet. Any account of consciousness which does not explain those properties is not an account of consciousness, at least in the sense identified by Chalmers' hard problem. These other accounts may be fruitful and scientifically valuable (as solutions to easy problems); they just are not solving the hard problem.

Some critics, on the other hand, argue that our responses about consciousness are themselves problematic, and so should not be taken as fixing a topic for investigation in this strong sense. Dennett, for example, argues that our responses about consciousness and qualia are confused, and so lead to confused philosophical theory (Dennett, 2018a). In an exchange with Galen Strawson in the *New York Review of Books,* Dennett writes:

> We [the 'consciousness deniers'] say that there isn't any conscious experience in the sense that Strawson insists upon. We say consciousness seems (to many who reflect upon the point) to involve being "directly acquainted," as Strawson puts it, with some fundamental properties ("qualia"), but this is an illusion, a philosopher's illusion. (Dennett, 2018b)

That is, Dennett challenges the topic-fixing responses, arguing that we are mistaken about our own conscious experience. This mistake, if this mistaken conception constrains what counts as an adequate theory of consciousness, likely will lead us towards false and inadequate accounts of

consciousness, and away from approaches that are more promising. It is not my task here to

wade into such disputes, but rather to note the structure of the debate. Cases are used to elicit

responses which, in the material mode identified by Machery, are used to make claims about the

meanings of terms. These claims are then used to indirectly constrain philosophical and scientific

theorizing.

2.2 Etiological Approaches

Etiological approaches use cases in order to get to the factors underlying the response,

and draw inferences on the basis of those factors rather than inferring directly that the content of

the response is true. The most obvious example is using our responses to cases to draw

inferences about the psychology of respondents. Following Sytsma and Livengood's taxonomy

of experimental philosophy (Sytsma & Livengood, 2016), I will call this the *cognitive* use of

cases. For example, trolley problems are often used in moral pyschology to probe what our moral

judgments are sensitive to (Mikhail, 2011).

In experimental philosophy, Joshua Knobe's work on the side-effect effect is a

paradigmatic example. Knobe's initial experimental work looked at the relationship between

normative evaluation and judgments about intentional action (Knobe, 2003). For example,

participants read about a CEO's actions that either helped or harmed the environment. They were

then asked if the CEO intentionally helped (or harmed) the environment. While 82% of the

participants in the harm condition judged that the CEO acted intentionally, only 23% of the

participants in the helped condition judged it to be an intentional action. This is a surprising

result. Antecedently, we might have expected that judgments about whether or not an action was

intentional would be made *prior* to evaluation, since we often think that intentionality matters to

that evaluation (for example, murder is a more serious crime than manslaughter because the death was intentional). This effect, Knobe has argued, is not merely a peculiar feature of our folk concept of intentional action. Rather, it is indicative of the ways in which moral judgment has a pervasive influence on a range of judgments in folk psychology (specifically those involving pro-attitudes towards outcomes) (Pettit & Knobe, 2009).

This is not to reduce the etiological approach to 'mere psychology,' that is, work without philosophical import. Putting aside whether a sharp boundary can be drawn between the interests of philosophers and those of psychologists, claims about why we respond to cases in the way we do are often then used as part of straightforwardly philosophical arguments. Joshua Greene, for example, uses work on trolley problem cases to challenge deontological approaches to ethics, and in defense of utilitarianism (Greene, 2013, 2007). Looking at both our responses to cases, and fMRI data on brain activity while responding, Greene argues that characteristically deontological judgments are rooted in affective processing, while characteristically utilitarian judgments are more cognitive. This, he argues, gives us a debunking argument against evidence in favor of deontology, and so has implications for normative theory.

The cognitive approach, however, is not the only one falling within the 'etiology' category. Cases might also be used in what I will call the *consistency with folk theory* interpretation. The consistency interpretation is committed to two claims: (1) that responses to cases are one source of evidence for claims about folk theory, and (2) we should strive to maintain consistency with that folk theory. Following Knobe's definition of folk psychology (Knobe, 2006), I define a folk theory as a conceptual framework underlying the inferences agents make in the everyday contexts relevant to that theory. We regularly, for example, ascribe mental states and predict actions on the basis of those ascriptions using folk psychology, while we

predict the movement of physical objects using folk physics. Here I will identify two variations of the consistency with folk theory interpretation that differ with regard to why we ought to strive for consistency with folk theory.

First, we may care about consistency with folk theory because such consistency is an epistemic virtue of a successful theory. The virtues are characteristics of good theories, and we can evaluate theories based upon how well they live up to those virtues. Kuhn identifies five virtues of good scientific theories: accuracy, consistency, broadness of scope, simplicity, and fruitfulness (Kuhn, 1977). An accurate theory makes successful predictions. A consistent theory is both internally self-consistent, and consistent with other accepted theories. A theory which is broad in scope covers more than just the phenomena it was designed to explain. A simple theory unites many phenomena under fewer, rather than more, theoretical commitments. A theory is fruitful when it directs us towards new phenomena of interest.

Responses are used as evidence of what folk theory is committed to, as it is presumed that this folk theory is part of the explanation for *why* we offer that response. Philosophers may differ in how much they care. Some might agree that consistency matters, but think that consistency with scientific theories is more important than with folk theories, and so not be moved by the lack of consistency revealed by a case, while others might dismiss the relevance of folk theory altogether. Such conflicts are, in Kuhn's view, unresolvable with any kind of decision procedure. There are two principal barriers to devising such a design procedure. First, theorists differ in their determinations of the most important domains for the virtues to cover. Two theories, A and B might have similar predictive success, but each makes predictions about slightly different sets of phenomena. If I think that the observations predicted by A are more important, but you think those predicted by B are more important, we cannot appeal to the virtue

of accuracy to resolve our dispute. Instead, Kuhn argues, it comes down to differences between us as theorists and differences in what we value. Such considerations are not amenable to a clear decision procedure.

Second, theorists differ about the relative weights of the virtues. For example, Chomsky describes his work in linguistics as "Galilean Science" (Chomsky, 2002, pp. 98–100), in that it prioritizes simplicity and fruitfulness ahead of accuracy. Galileo, Chomsky argues, offered a theory that did not explain some seemingly obvious data points, such as why the rotation of the Earth does not fling us off into space. Chomsky saw his work on the Minimalist program in syntax as similarly yielding ground on coverage of the data. Both theories, however, were simpler in that they reduced the number of principles required to explain the data, and more fruitful in that they suggested more directions for future research. Once again, these differences come down to personal differences between theoreticians, rather than divergences from a single shared epistemic standard.

It is worth noting how this interpretation differs from using folk theory as a bridge to infer the truth of the response. One might argue that folk theory is generally reliable, and so individual judgments rooted in folk theory are also reliable. This reduces consistency with folk theory to accuracy. Yet, this kind of appeal to theoretical virtue is useful precisely in the case where we cannot so readily establish that any one of our theories is sufficiently reliable (and so we turn to other markers of a successful theory). Indeed, an advantage of the interpretation of the role folk theory plays that I have provided is that we can avoid the assumption that folk theory is reliable, particularly for the complex and esoteric cases that philosophers typically investigate, and instead treat consistency with folk theory as just one virtue among many of a successful theory

The second reason one might care about consistency with folk theory in the first place is because consistency with that theory may have practical effects we ought to be concerned with. Robinson & Darley (see also Carlsmith & Darley, 2008; Robinson & Darley, 2007) provide an example of a case where intuitions might be sufficiently important to play a decisive role in policy decisions. They argue that people across cultures (Carlsmith, 2006; Darley, Carlsmith, & Robinson, 2000; Robinson & Kurzban, 2007) have retributivist intuitions about punishment. Retributive intuitions are contrasted with consequentialist intuitions. A retributive intuition is one where punishment is justified on the basis of the match (or challenged on the basis of a mismatch) between a punishment and the deservingness of the perpetrator. A consequentialist intuition, by contrast, is one where the punishment is justified on the basis of the goods that result from that punishment. For example, a retributivist might punish a thief on the grounds that the thief deserves a punishment which fits the crime (and thus, not something either too harsh or too lenient), while the consequentialist might punish the thief to deter future crimes, or to incapacitate the thief and prevent future theft from that individual.

Adopting a justice system which is contrary to these widely shared retributive intuitions, Robinson & Darley argue, would have substantial negative consequences. It would undermine the stigmatization assigned to lawbreakers, a stigmatization which plays an important role in voluntary compliance to the law. It would also undermine faith and confidence in the legal system, which in turn promotes vigilantism and subversion of the system. Finally, it would undermine the moral credibility of the legal system, making it a less trustworthy guide to permissible behavior, particularly in borderline cases. Since changing these intuitions is difficult, and liable to cause even greater harm (by violating other moral principles, such as those prohibiting coercive indoctrination), we have reason to adopt public policies which respect these

intuitions. Indeed, given the severity of the consequences for contravening our intuitions, consistency with the intuitions might be the decisive reason in favor of a retributive justice system.

Robinson & Darley's argument can plausibly be reconstructed in terms of folk theory, rather than intuitions. After all, it is not merely that a consequentialist system of justice would violate *only* our intuitions; the problem is that it contravenes strongly held convictions that manifest in intuition, and plausibly, behavior. Their argument can then be schematized as follows:

1. People (fixed relative to the population in question) have strong intuitions about *x*.

2. Strong intuitions about *x* indicate that according to the folk theory shared by those people, *p*

3. The costs of contravening or changing that folk theory exceeds the cost of adopting policies which are consistent with it.

4. Policies should be adopted which maximize the benefits relative to the costs of the policy relative to competing policies.

5. Therefore, we should adopt policies consistent with the folk theory.

This case exemplifies how the etiological approach differs from the reliabilist approach. Our responses to cases about punishment are used to draw inferences about the folk theory underlying them. Yet, we need never draw any conclusions about the truth of those responses. Instead, they are used as evidence for what people think, and what they think can play one of several roles in philosophical argumentation,

2.3 Non-evidential Approaches

In this section, I turn to non-evidential uses of cases, where the cases are not used as evidence for any claims, whether about their causes or their contents. Machery has already suggested two types of non-evidential uses: the provocative and illustrative uses. There are a number of essentially pedagogical purposes to which a thought experiment may be put. One is that a thought experiment makes an argument more concrete, tying it to particulars. The design of a thought experiment may also help the audience better understand the particular property under investigation by distinguishing it from other, related, properties. The consistency between our intuitive responses to different cases may also help us to better understand the attractiveness of a theory without thereby being evidence for it. While this may simply be an instance of rhetorical persuasion, leading an audience to seeing the attactiveness of a position may also lead to a more charitable interpretation of the argument (rather than operating in place of it).

Consider, for example, how one might read Thomson's violinist case as an illustrative use (Thomson, 1971). Thomson asks us to imagine that one wakes up only to find that a famous violinist has been attached to you through a complex machine. You are told that you need to keep the violinist attached to you for nine months, as this machine keeps the violinist alive, and if detached, the violinist will die. You did not consent to having the violinist hooked up to you in this way. Are you obligated, Thomson asks, to keep the violinist hooked up for those nine months? Thomson answers "I imagine you would regard this as outrageous, which suggests that something really is wrong with that plausible-sounding argument I mentioned a moment ago [that since a fetus has a right to life, abortion is immoral]" (Thomson, 1971, p. 49).

This term "suggests" is slippery. It is sometimes used evidentially; in this case, it would indicate that our responses here provide at least some evidence against the view that there is a

moral obligation not to have an abortion. Alternatively, however, it might be read as providing a *clue* as to where arguments might be marshalled. By a "clue," I mean that we might accept the counter-factual that, if our responses are reliable, the thought experiment shows that a right to life does not guarantee a right to care. Yet, one may not grant the antecedent of the conditional, and instead see if the consequent can be defended by other means. In this case, the thought experiment is non-evidential; it illustrates to the reader the central properties under consideration (right to life, and right to care from others) in a way which highlights and motivates providing arguments on these properties. It is, however, those arguments which carry the evidential burden. The role of the clue is two-fold. It suggests where the development of arguments may be promising (in the same way that an intuition in the sciences might suggest questions to explore experimentally), and it suggests to the audience why those arguments might be relevant to the topic.

The illustrative power of cases, may explain why thought experiments have such lasting power in philosophical discourse. If thought experiments make arguments, and the motivations for them, easier to understand, then we should expect those thought experiments to be more memorable than the more detailed arguments. That thought experiments play an outsized role in our understanding and memory of a debate does not entail that those thought experiments played an outsized role in the evidence pushing that debate forward.

An alternative non-evidential approach is the *common grounds* interpretation. On this approach (Cappelen, 2012; Chalmers, 2013; Maynes, 2017), cases and our responses to them are primarily used in a dialectical role. Stalnaker notes that a speaker "may presuppose any proposition that he finds it convenient to assume for the purposes of the conversation, provided that he is prepared to assume that his audience assumes it along with him" (Stalnaker, 1999, p.

84). That is, two assumptions are required to put a proposition in the common ground: the proposition itself must be assumed, and so too must it be assumed that one's interlocutors also assume the proposition. In philosophy, premises are assumed for a variety of reasons, whether for the sake of argument, because one presumes it has already been established elsewhere, or because one thinks it is a basic enough premise that it will be widely shared. Such premises are part of the common ground in that dialectical context.

Placing a proposition in the common ground is a dialectical move, not an epistemic one, even if that proposition goes on to play an evidential role. The proposition in the common ground may be inferentially justifiable, but that potential justification plays no role in the dialectical context in which that proposition is used (see also Chalmers, 2013). This may be because the proposition seems so obvious that no justification is demanded, that the reasons the proposition is justified are contentious even if its status as justified is not, or it may be that the proposition is so fundamental that it is unclear how that justification could be offered even though interlocutors share a commitment to it. On the common ground reading, one role for cases is to help make clear what we are already committed to, even if we are not explicitly aware of it.

At times, Searle describes his Chinese Room thought experiment in this way, as an illustration of the assumptions in the common ground. After presenting the Chinese Room argument, which he describes as a "common sense" (rather than as a theoretical) objection (Searle, 1992, p. 45), Searle considers a range of possible replies from defenders of the computational picture. The first is the "Systems Reply," or the argument that while the Chinese Room itself does not understand Chinese, if that room were embedded into a much larger, more complex system, we would agree that the system understands. In reply, Searle contends that even

if the individual in the Chinese Room internalized the whole system, he would still not understand. He continues:

> Actually, I feel somewhat embarrassed to give even this answer to the systems theory because the theory seems to me so unplausible to start with. … It is not easy for me to imagine how someone who was not in the grip of an ideology would find the idea at all plausible (Searle, 1980, p. 419).

Searle then provides further examples designed to dislodge this ideologically driven (or theoretically driven) response. Indeed, he closes his paper with a discussion of the illusions that have led people to think that Strong AI is true, with the aim of dispelling the seeming behind this program (Searle, 1980, p. 423). This "dispelling" language might be read epistemically, namely that the case provides a *reason* to reject the seeming. If, however, we are to suppose that champions of Strong AI are so deep in the "grip of an ideology," it is unlikely that a thought experiment would provide sufficient reason. Instead, Searle might be read as using the thought experiment illustratively, and that "dispelling the seeming" is best understood psychologically. The thought experiment, in illustrating the (assumed) absurdity of the consequences of Strong AI, is designed to help Strong AI advocates see that their views contradict the assumptions about mindedness they are already (and pre-theoretically) committed to.

For the idea that simple machines, such as thermostats, can have beliefs, Searle is even more direct about the our starting assumptions. "The study of the mind starts with such facts as that humans have beliefs, while thermostats, telephones, and adding machines don't," Searle remarks, "if you get a theory that denies this point you have produced a counterexample to the theory and the theory is false" (Searle, 1980, p. 420). Not only do we start with these facts, but

Searle does not psychologize the starting points. It is not that we start with responses and infer these facts; these facts are the bare assumptions we begin inquiry with.

What of more esoteric and complex cases, like the Chinese Room? Though Searle describes the case as a common sense objection in *The Rediscovery of Mind,* he does not always do so. In an exchange with Steven Pinker published in the *New York Review of Books,* Searle writes that "in over twenty years of debating these issues, I have never relied on common sense. I appeal to a logical distinction between the syntax of the implemented program and the semantics of actual human understanding, and the [Chinese Room thought experiment] is designed to illustrate the distinction between the syntax and the semantics" (Searle, 2002). Here, Searle suggests that his objection can be stated without appeal to the trappings of the Chinese Room at all. It is that strictly syntactic systems cannot, by their nature, give rise to semantic properties. Since Strong AI is built on the presumption that syntactic systems can do so, Strong AI is false.

At first glance, this might seem an uncharitable reading of Searle's argument. After all, this claim is precisely what is at issue between Searle and his opponents. Searle knows that Strong AI defenders would not grant him the assumption that syntax is insufficient for semantics, so it cannot be in the common ground. The key word in Searle's remarks above, however, is that the Chinese Room illustrates this claim. On the common ground reading, one role for thought experiments is to help make clear what we are already committed to, even if we are not explicitly aware of it. This is why Searle sees his arguments as dispelling the seeming behind Strong AI. Searles takes it that even defenders of Strong AI are, in fact, already committed to the view that the Chinese Room does not understand. This commitment, however, is obscured by a number of theoretical commitments brought on by work within the computationalist program. The power of

a thought experiment is that its simplified structure helps to bring out tacit commitments by moving us from language laden with theoretical baggage.

A common grounds approach is suggested by Cappelen (2012), and considered by Machery, who rejects it, asking:

> Why are we warranted in assuming, e.g., that it is permissible to cause someone's death in the situation described by the switch version of the trolley case? It can't simply be because philosophers view them as belonging to the common ground. Assuming that *p* does not make one warranted to assume that *p*. I wish it was that easy! (Machery, 2017, p. 178)

This focus on *warrant* for the common ground is suggested by Cappelen, who in describing the the interpretation of "intuitive" as "pre-theoretic" (i.e., belonging to the common ground), writes that "the speaker is saying that *p* is a judgment that has been or can be justified without taking a stand on the question under discussion" (Cappelen, 2012, p. 81). On the proposal offered here, however, the MoC is not intended to provide warrant for taking a proposition as part of the common ground. Instead, cases play a dialectical role in helping one's interlocutors see that they share that assumption. This may, certainly, fail, and if so, the use of the case is defective. It may also be that my interlocutors and I share assumptions that are limited and parochial, and I agree with Machery that experimental philosophy has a crucial role to play in figuring out when this is the case. My point here is more limited: it is that the common grounds interpretation of the MoC helps us to better understand one important use of cases, and how those cases might be used without relying on implicit appeals to the reliability of intuition or any other type of response.

3. Taking Stock

Let us close by stepping back and taking stock of the methods discussed above. First, I introduced three categories:

Reliabilist Interpretations - responses are taken to be reliable indicators of the truth of their contents.

Etiological Interpretations - responses are used to draw inferences about the factors that are causally responsible for the production of the response.

Non-Evidential Interpretations - responses are used for pedagogical or dialectical value, not to establish the truth of any claim.

Within the reliabilist interpretations, I first distinguished a range of possible interpretations based upon the degree of credence the response gives its content. These interpretations can range from infallible intuitions to merely a *prima facie* justification. This range of interpretations suggests that the Strong MoC is too limiting, as if our responses to cases play only a *prima facie* role in justifying claims, we cannot infer which theory is correct solely on the basis of consistency with those responses. I then argued that Alvin Goldman's *topic-fixing* approach is a type of reliabilist interpretation which uses cases to identify facts about meanings, then uses those claims about meanings to indirectly constrain philosophical and scientific theorizing. That is, further work on *x* should be consistent with certain crucial features of the meaning of "*x*" *if* it is to count as an investigation of *x*.

Within the etiological interpretations, I identified two methods. The first, the *cognitive* approach uses our responses to cases as data points in drawing inferences about the psychology of the responder. The second, the *consistency with folk theory* approach uses our responses to

draw inferences about the content of folk theory. I then identified two roles folk theory might play in philosophical arguments; first, if one values consistency with folk theory as a theoretical virtue, and second, if one values consistency with folk theory based on the consequences of that consistency.

Finally, I identified three non-evidential interpretations. Following Machery, on the *provocative* and *illustrative* use of cases, the case is used for its pedagogical value in helping an audience to understand and reflect upon ideas. On the *common grounds* approach, cases are used to help the audience recognize an antecedent commitment, and so accept a proposition as being part of the common ground. In each of these interpretations, the case plays a role in eliciting a response that plays a dialectical or pedagogical, rather than evidential role.

Thus, I conclude, there is no *method* of cases, there are multiple *methods* of cases. There are two principal pay-offs to this result. First, it challenges Machery's uniformity thesis. Cases are widely used in philosophy, but cases are tools that can be used in a variety of ways. In identifying this range of uses, and by showing that they are plausible interpretations of paradigmatic appeals to cases, I have shown that we have reason to prefer methodological heterogeneity. Beyond the empirical question of whether or not a hetereogenous interpretation can be supported with examples, Machery contends that we should expect homogeneity as an explanation of philosophers' ability to communicate successfully about cases. Yet, I take one lesson of X-Phi's focus on the MoC to be that philosophers have communicated less effectively than we might have thought in our uses of cases. Distinguishing these various methods may reveal distinctions that philosophers have hitherto missed in understanding the work of others. Indeed, this may be one explanation for why cases play such a prominent role in the proliferation

of philosophical arguments - the contours of their original use are lost, but a memorable case is easily interpreted as having evidential value even if not originally intended to play such a role.

Second, these interpretations help stock our methodological toolbox. This not only helps us be more charitable towards the arguments in question, but clarifies the philosophical significance of X-Phi more generally, and specific findings in particular. Debates over X-Phi's philosophical significance often hinge on a Method of Cases reading of traditional philosophy, so much so that some critics take a rejection of this reading as grounds for a rejection of X-Phi. Developing a better suite of interpretive frameworks for understanding philosophical methodology will help us to give a more nuanced and subtle reading of the roles cases plays in philosophical argument. This better positions us to recognize how these arguments work, what experimental work would be valuable, and what the significance of that work is. This helps us to replace a dead-end question (is X-Phi philosophically relevant?) with more useful ones - when is it relevant, and in what way?

References

Alexander, J. (2012). *Experimental Philosophy: An Introduction*. Malden, MA: Polity Press.

Bealer, G. (1998). Intuition and the Autonomy of Philosophy. In M. DePaul & W. Ramsey (Eds.), (pp. 201–40). Rowan & Littlefield Publishers.

Block, N. (2007a). Troubles with Functionalism. In *Consciousness, Function, and Representation: Collected Papers* (Vol. 1, pp. 63–102). Cambridge, Mass.: The MIT Press.

Block, N. (2007b). What intuitions about homunculi don't show. In *Consciousness, Function, and Representation: Collected Papers* (Vol. 1, pp. 103–108). Cambridge, Mass.: The MIT Press.

Cappelen, H. (2012). *Philosophy Without Intuitions*. Oxford University Press.

Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology*, *42*(4), 437–451.

Carlsmith, K. M., & Darley, J. M. (2008). Psychological Aspects of Retributive Justice Other Kinds of Justice. In M. Zanna (Ed.), (Vol. 40, pp. 193–236). San Diego: Elsevier.

Chalmers, D. J. (1997). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

Chalmers, D. J. (2003). Consciousness and its Place in Nature. In S. Stich & T. Warfield (Eds.), *Blackwell Guide to the Philosophy of Mind*. Blackwell.

Chalmers, D. J. (2013). Intuitions in Philosophy : A Minimal Defense, (1992), 1–11.

Chomsky, N. (2002). *On Nature and Language*. Cambridge University Press.

Cummins, R. (1998). Reflection on Reflective Equilibrium. In M. R. DePaul & W. Ramsey (Eds.), (pp. 113–128). Rowman \& Littlefield Publishers, Inc.

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and human behavior*, *24*(6), 659–83.

Dennett, D. (2018a). Facing up to the hard question of consciousness. *Phil. Trans. R. Soc. B*, *373*(1755), 20170342.

Dennett, D. (2018b). "Magic, Illusions, and Zombies": An Exchange. *The New York Review of Books*.

Devitt, M. (2006). Intuitions. In V. G. Pin, J. I. Galparsoro, & G. Arrizabalaga (Eds.), (pp. 169–176). Universidad del Pais Vasco.

Goldman, A. I. (2007). Philosophical Intuitions: Their Target, Their Source, and Their Epistemic Status. *Grazer Philosophische Studien*, *74*, 1–26.

Greene, J. (2013). *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. New York: Penguin Press HC, The.

Greene, J. D. (2007). The Secret Joke of Kant's Soul. In W. Sinnott-Armstrong (Ed.), (pp. 35–80). The MIT Press.

Knobe, J. (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, *16*, 309–323.

Knobe, J. (2006). The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology. *Philosophical Studies*, *130*, 203–231.

Kornblith, H. (1998). The Role of Intuition in Philosophical Inquiry: An Account with no Unnatural Ingredients. In M. DePaul & W. Ramsey (Eds.), (pp. 129–141). Rowman & Littlefield Publishers.

Kuhn, T. S. (1977). Objectivity, Value Judgment, and Theory Choice. In *The Essential Tension* (pp. 320–39). University of Chicago Press.

Ludwig, K. (2007). The Epistemology of Thought Experiments: First Person versus Third Person Approaches. *Midwest Studies in Philosophy*, *XXXI*, 128–59.

Lycan, W. (1988). *Judgment and Justification*. Cambridge University Press.

Machery, E. (2017). *Philosophy Within Its Proper Bounds*. New York, NY: Oxford University Press.

Mallon, R., Machery, E., Nichols, S., & Stich, S. (2009). Against Arguments from Reference. *Philosophy and Phenomenological Research*, *LXXIX*(2), 332–356.

Maslow, A. (1966). *The Psychology of Science*. New York: Harper and Row.

Maynes, J. (2017). On the stakes of experimental philosophy. *Teorema: Revista internacional de filosofía*, *36*(3), 45–60.

Mikhail, J. (2011). *The Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge University Press.

Nahmias, E., Morris, S., Nadelhoffer, & Turner, son. (2005). Surveying Freedom: Folk Intuitions about free will and moral responsibility. *Philosophical Psychology*, *18*(5), 561–584.

O'Neill, E., & Machery, E. (2014). Introduction: Experimental Philosophy: What Is It Good For? In *Current Controversies in Experimental Philosophy* (pp. vii–xxix). New York and London: Routledge.

Pettit, D., & Knobe, J. (2009). The Pervasive Impact of Moral Judgment. *Mind & Language*, *24*(5), 586–604.

Putnam, H. (1975). The Nature of Mental States. In *Mind, Language, and Reality*. Cambridge: Cambridge University Press.

Robinson, P. H., & Darley, J. M. (2007). Intuitions of Justice: Implications for Criminal Law and Justice Policy. *Southern California Law Review*, *81*(1).

Robinson, P. H., & Kurzban, R. (2007). Concordance & Conflict in Intuitions of Justice. *University of Pennsylvania Law School, Public Law Working Paper No. 06-38, Minnesota Law Review*, *91*(6).

Searle, J. (2002). "Words and Rules": An Exchange. *The New York Review of Books*.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(03), 417–424.

Searle, J. R. (1992). *The Rediscovery of the Mind* (Massachusetts Institute of Technology edition.). Cambridge, Mass: A Bradford Book.

Stalnaker, R. C. (1999). Assertion. In (pp. 78–95). Oxford University Press.

Sytsma, J., & Livengood, J. (2016). *The Theory and Practice of Experimental Philosophy*. Broadview Press.

Thomson, J. J. (1971). A Defense of Abortion. *Philosophy and Public Affairs.*, *1*(1), 47–66.

Williamson, T. (2011). Philosophical Expertise and the Burden of Proof. *Metaphilosophy*, *42*(3), 215–229.