

## **Pragmatic Experimental Philosophy**

**Justin C. Fisher – Southern Methodist University**

**Abstract:** This paper considers three package deals combining views in Philosophy of Mind, Meta-Philosophy, and Experimental Philosophy. The most familiar of these packages gives center-stage to pumping intuitions about fanciful cases, but that package involves problematic commitments both to a controversial descriptivist theory of reference and to intuitions that ‘negative’ experimental philosophers have shown to be suspiciously variable and context-sensitive. In light of these difficulties, it would be good for future-minded experimental philosophers to align themselves with a different package deal. This paper suggests two alternatives. Experimentalists might help fans of “Naturalized” approaches discover what natural kinds have been playing an appropriate role in causing us to use concepts as we do. Or, better still, experimentalists might instead help pragmatists and teleo-semanticists discover how our concept usage regularly yields beneficial outcomes, so that we can then craft philosophical analyses that will enable us to yield such beneficial outcomes more consistently. Using *free will* and *explanation* as instructive examples, this paper offers concrete guidance and suggestions for how experimental philosophers can pursue new positive projects that will be both pragmatically and philosophically useful.

### **1. Introduction**

This paper explores related issues in three areas of philosophy that don’t interact nearly as much as I think they should. First, there are people working in *philosophy of mind and philosophy of language* who develop general theories about what our words and concepts<sup>1</sup> refer to, or what things they would be correctly applicable to. Second, many analytic philosophers spend at least some time doing *meta-philosophy*, talking about the methods we might use to generate analyses of philosophical concepts—i.e., to generate clear articulations of what exactly those concepts refer to or are correctly applicable to. And third, there is a growing movement of *experimental philosophers* who think philosophers shouldn’t just sit and reflect in armchairs, but should instead actively seek empirical evidence to test and support our views.

---

<sup>1</sup> I think of concepts as mental particulars that play a folder-like information-coordinating role in cognition. In this, I follow a rich tradition in cognitive psychology and the philosophy of cognitive science. For a good introduction to this tradition see Lawrence & Margolis (1999), and for a view of concepts very similar to my own see Millikan (1998, 2000). This tradition may be contrasted against an equally rich philosophical tradition that takes concepts to be abstract entities.

There are strong connections between these three areas. Any view about the correct applicability of concepts will have implications both in meta-philosophy (regarding what sort of methodology philosophers might use to articulate what philosophical concepts are correctly applicable to) and in experimental philosophy (regarding what sorts of empirical evidence we would need to seek in order to decide which proposed analyses are correct). Hence, views in any of these three areas are naturally packaged with commitments in the others. In previous work (Fisher 2006, forthcoming), I have considered three plausible ‘package deals’ combining methodologies defended by meta-philosophers with theories of reference defended by philosophers of mind and language. The primary aim of the present paper is to extend this work to consider the sorts of *experimental* methodologies that would best complement each of these package deals.

My overview of these potential positions will, I hope, be fair. But it will not be impartial. For I have a special interest in developing and exploring pragmatic philosophical and experimental methodologies that couple nicely with *teleosemantic* theories of reference. These pragmatic methodologies have been under-represented in both meta-philosophy and experimental philosophy. This paper is part of an ongoing project to draw attention to the relations between these three areas, and in particular to attract adherents, including experimental philosophers, to the package deal that I favor. Regardless of whether I manage to make new converts to my teleo/pragmatic approach, this paper will highlight important connections between our theories of reference and our philosophical methodologies, and it will help to make clear the ways in which experimental findings about concept usage may or may not be philosophically relevant.

## 2. Three Package Deals

The rows in Table 1 summarize what I see as the three most viable package deals. I'll talk about each in turn.

Theory of Reference	Philosophical Methodology	Empirical Methodology
<b>Descriptive theories</b> hold that a concept is correctly applicable to whatever fits the description we have subconsciously associated with that concept.	<b>Intuitive Conceptual Analysis</b> seeks an explication that best fits our intuitions involving a concept (which hopefully will reflect the descriptions hidden inside us).	<b>Traditional Experimental Philosophy</b> gives carefully designed prompts to various subjects and then observes their intuitive responses.
<b>Causal/Informational theories</b> hold that a concept is correctly applicable to whatever has played an appropriate role in causing its usage.	<b>Naturalized Philosophy</b> identifies some clear instances of a concept (typically instances that caused us to use that concept) and then seeks a natural kind that includes those instances.	<b>Naturalized Experimental Philosophy</b> tests to see which things play an appropriate role in causing subjects' usage of a concept and uses science to discover what natural kind(s) those things belong to.
<b>Teleosemantic/Pragmatic theories</b> hold that a concept is correctly applicable to whatever would best sustain existing patterns in its beneficial usage.	<b>Pragmatic Conceptual Analysis</b> seeks an explication that will best preserve the patterns of beneficial usage for a given concept.	<b>Pragmatic Experimental Philosophy</b> tests to see when and how subjects regularly benefit from employing a concept.

**Table 1.** Three package deals (rows) each combine a theory of reference, a philosophical methodology, and an accompanying empirical methodology.

A great deal of work in analytic philosophy employs the methodology I call *Intuitive*

*Conceptual Analysis*. One staunch defender of this methodology is Frank Jackson:

[H]ow should we identify our ordinary conception? The only possible answer, I think, is by appeal to what seems to us to be most obvious and central about free action, determinism, belief, or whatever, as revealed by our intuitions about possible cases. [...] Intuitions about how various cases, including various merely possible cases, are correctly described in terms of free action [etc...] are precisely what reveal our ordinary conceptions of free action [etc...] (Jackson 1998, p. 31).

This intuition-pumping methodology would make sense if we still held old-fashioned descriptive theories of reference,<sup>2</sup> as these theories say that reference is determined by the descriptions we

<sup>2</sup> Descriptivist views include Frege (1892), Russell (1905), Strawson (1950).

tacitly associate with our concepts, and intuitions might plausibly reflect those tacit descriptions. But, most people working in philosophy of mind and language have abandoned descriptive theories of reference<sup>3</sup> and moved on to either causal/informational<sup>4</sup> or teleo/pragmatic theories.<sup>5</sup> On these theories, reference is determined in large part by patterns of causal interaction between us and things in the world.

The mass exodus from descriptivism bodes poorly for Intuitive Conceptual Analysis. Granted, some intuitions may track the semantically relevant external causal relations. Indeed, insofar as our intuitions are shaped by practical experience (via either learning or evolution), it may be especially likely that we'll end up having quite good intuitions regarding ordinary practically relevant cases. However, Intuitive Conceptual Analysis has been characterized by its consideration of cases that are far from ordinary: cases like Swampman, the Chinese Room, TrueTemp, or Frankfurt's neuroscientist. For extraordinary cases like this, there is no reason to think intuitions will do especially well at tracking truths about the referents we have latched onto via the sorts of causal interaction that are deemed semantically relevant by most contemporary semantic theories. Furthermore we now know that folk intuitions *often do fail* when extrapolated to extraordinary cases: despite intuitions to the contrary, solid objects aren't impermeable at small scales, water isn't an indivisible element, species don't have essences that are immutable over evolutionary time-scales, and space-time isn't Cartesian at small scales or high speeds. It would take great hubris to expect folk philosophical intuitions to fare better than folk intuitions

---

<sup>3</sup> The exodus from descriptivism was spurred in part by classic examples by Kripke (1972), Putnam (1973) and Burge (1979). These include examples where people don't have in mind descriptions sufficient to pick out their referents uniquely (e.g., Kripke's Feynman case, Putnam's Twin Earth), and examples where the descriptions people do have in mind don't even fit their actual referents (e.g., Kripke's Gödel/Schmidt case, Burge's arthritis case).

<sup>4</sup> Causal/Informational approaches include Kripke (1972), Evans (1973), Boyd (1988), Fodor (1990), Rupert (1999).

<sup>5</sup> Teleo/Pragmatic semantic theories include James (1906), Ramsey (1927), Millikan (1984), Appiah (1986), Papineau (1987), Dretske (1988), Whyte (1990), and Blackburn (2005).

surrounding other concepts.<sup>6</sup> Descriptivism offered one potential basis for such hubris: our intuitions would be trustworthy if they reflected tacit descriptions that determined reference. But once we reject descriptivism, we lose our grounds for optimism about the reliability of intuitions regarding extraordinary cases, and hence we lose our grounds for trusting Intuitive Conceptual Analysis.

Even while our best semantic theories give us strong reason to doubt our intuitions, a great deal of work in analytic philosophy still proceeds by pumping intuitions. This is an unstable situation. Sooner or later some sort of correction will need to be made to resolve it: either philosophers of mind and language will need to reincarnate descriptivism,<sup>7</sup> or we'll all need to adopt new methodologies that better fit our state-of-the-art theories of reference.

Two alternative methodologies are especially plausible. First, corresponding to causal/informational theories of reference (which hold that our concepts refer to the (kinds of) things that play an appropriate role in causing our use of those concepts), there is what is often called "*Naturalized philosophy*". One strong advocate<sup>8</sup> of this methodology is Hilary Kornblith:

We begin, often enough, with obvious cases, even if we do not yet understand what provides the theoretical unity to the kind we wish to examine. *Understanding what that theoretical unity is* is the object of our study, and it is to be found by careful examination of the phenomenon, that is, something outside of us, not our concept of the phenomenon, something inside of us. (Kornblith 2002, pp. 10-11.)

Second, corresponding to teleo/pragmatic theories of reference, there is the methodology that I have called *Pragmatic Conceptual Analysis*, a methodology that focuses not on what has been *causing* us to apply our concepts, but instead upon the *beneficial effects* that our concept usage

---

<sup>6</sup> In a moment we'll see that "nay-saying" experimental philosophers have further called into question the reliability of some particular philosophical intuitions by showing that these intuitions vary across people and/or contexts.

<sup>7</sup> Some contemporary philosophers intend to do just that, including Lewis (1984) and Chalmers (2002, 2004).

<sup>8</sup> Philosophers have proposed "Naturalized" analyses of many philosophical concepts, including knowledge (Quine 1969, Kornblith 2002), emotion (Griffiths 1997), color (e.g., Hilbert 1987), consciousness (Dennett 1991), and moral goodness (e.g., Boyd 1988).

has regularly caused. The epistemologist Sally Haslanger offers a clear articulation of how one might apply Pragmatic Conceptual Analysis to the concept of knowledge:<sup>9</sup>

[T]he best way of going about a project of normative epistemology is first to consider what the point is in having a concept of knowledge: what work does it, or (better) could it, do for us? And second, to consider what concept would best accomplish this work. (Haslanger 1999, pg 467).

I have argued elsewhere<sup>10</sup> that Pragmatic Conceptual Analysis not only fits well with popular teleosemantic theories; it also provides explications that are practically worth adopting; it is a formal version of positive reinforcement learning processes that we subconsciously employ all the time; it alone provides explications that enable us to give satisfying explanations of people's use of concepts; and it is supported by the most plausible reading of the Principle of Charity.

A brief aside: even if you reject teleosemantics, Pragmatic Conceptual Analysis may still be worth employing as a way to identify explications that we'll have practical reason to adopt. For example, a pragmatic analysis of "knowledge" would tell us what things we'd need to count as knowledge in order best to continue the sorts of useful work we've been doing by distinguishing "knowledge" from "mere belief". Even if you doubt that such an analysis would capture what "knowledge" really means – i.e., even if you think this analysis would describe some other thing, knowledge\*, rather than actual knowledge – it can still be very useful to have achieved a good characterization of knowledge\* and it will likely be worth coining a term for knowledge\* or perhaps even stipulating that, henceforward, we'll use the word "knowledge" to mean knowledge\*. As a teleosemanticist, my view is that "knowledge" already means knowledge\*, so explicating "knowledge" as knowledge\* is mere clarification rather than revisionary stipulation. But even if you reject this teleosemantic claim, you should still recognize that knowledge\* is

---

<sup>9</sup> Other articulations of Pragmatic Conceptual Analysis are given by Edward Craig (1986, 1990), Jim Woodward (2003), and my own Fisher (2006).

<sup>10</sup> See Fisher (2006; forthcoming; in prep.).

worth identifying and coining *some* term for. And hence you should recognize that Pragmatic Conceptual Analysis is a methodology worth exploring.

However, the point of the present paper is not to sing the praises of Pragmatic Conceptual Analysis. Instead our aim is to explore the *experimental* methodologies that would best correspond to this and to the other ‘package deals’ mentioned above.

### 3. Cautionary Optimists, Nay-Sayers, and Curiosity Seekers

Most work in experimental philosophy (or ‘X-Phi’) has involved pumping intuitions from various experimental subjects. Most experimental philosophers have fallen into at least one of the following three sorts of intuition-pumpers.

**Curiosity Seekers** attempt to elucidate the cognitive processing underlying subjects’ intuitive usage of philosophical concepts (e.g., knowledge or freedom), but do not attempt to draw conclusions about the nature of the referents of those concepts (e.g., about the nature of knowledge or of freedom).<sup>11</sup>

**Nay-Sayers** attempt to undercut philosophers’ use of intuitions to support philosophical views. Some nay-sayers seek only to show that many people do not have the intuitions philosophers have claimed them to have. Other nay-sayers attempt to show that we should generally expect intuitions to be a poor guide because they are too variable, context-sensitive, and/or culture-relative.<sup>12</sup>

**Cautionary Optimists** hold that intuition-pumping is a good way to do philosophy but warn that we need to be very careful whose intuitions we pump and how we pump them.<sup>13</sup>

Of course, some particular philosophers might fall in different categories at different times in their careers or even at different points in the same paper (e.g., exploring some results just for curiosity’s sake while drawing positive or negative conclusions from others). I take these three

---

<sup>11</sup> See, e.g., Nichols & Knobe (2008).

<sup>12</sup> Nay-sayers include Weinberg, et al (2001), Machery et al (2004), Weinberg (2007), Swain, et al (2008), Mallon et al (2009), and Alexander et al (forthcoming) which is appropriately entitled “Accentuate the Negative.”

<sup>13</sup> Some advocates of intuitive conceptual analysis have indicated openness to having experimentalists gather intuitions, including Lewis (1972) and Jackson (1998). In their “Experimental Philosophy Manifesto” Nichols & Knobe (2008) emphasize the continuity of experimental philosophy with traditional philosophical use of intuition. Many experimental philosophers have drawn positive philosophical conclusions from the intuitions they evoke, including Nahmias et al (2006), Griffiths & Stotz (2008), and Roskies & Nichols (2008).

categories to be quite exhaustive of possible justifications for intuition pumping. Anyone who thinks intuition-pumping is worth doing must think either that it can yield positive philosophical conclusions about the referents of the concepts in question (cautionary optimists), that it can undercut some philosophers' attempts to yield such conclusions (nay-sayers), and/or that it is worth doing for other reasons besides supporting or undercutting arguments for such conclusions (curiosity seekers).

Nichols and Knobe (2008) argue that there is a long tradition of curiosity-seeking in philosophy, including Plato, Aristotle, Spinoza, Hume and Nietzsche. Curiosity about human cognition is admirable, and I don't mind construing "philosophy" broadly enough to count this as a sort of philosophy. However, it's worth emphasizing that there are deeper questions to ask about (for example) knowledge and freedom, beyond asking how ordinary people think about knowledge and freedom. Epistemologists would like to know, for example, whether knowledge *actually does* require absolute certainty, regardless of what people think. And action theorists would like to know whether any of our actions *actually are* free, regardless of what people think. When purely curiosity-seeking experimental philosophers avoid these deeper questions, they thereby relegate themselves to the margins of substantive philosophical debates.

Unlike mere Curiosity Seekers, Nay-Sayers hold that empirical findings about our intuitions do have at least some potential bearing upon these substantive philosophical debates. However, their findings are uniformly negative: people often don't have the intuitions philosophers have claimed them to have, and/or the intuitions that people do have seem to be too variable, too context-sensitive, or too heavily influenced by culture, gender, personality type, or socio-economic class to provide a good guide to philosophical matters.

I think Nay-Saying experimental philosophers are like dragon slayers. So long as we're tormented by dragons, it's great to have people around to slay them. Similarly, as long as we're tormented by blind followers of intuition, it's great to have Nay-Sayers around to set them straight. However, dragon-slayers face a lose-lose scenario: either they'll fail to slay all the dragons, or they'll succeed and drive themselves out of business. Nay-Sayers face the same two unfortunate options. They might fail to make mainstream philosophers abandon their appeals to intuition, or they might succeed, leaving themselves no more intuition-mongers to hunt.<sup>14</sup> Nay-Sayers have undertaken an entirely negative project: showing that intuitions are too strange or variable to play the positive role that other philosophers have wanted them to play. But it would be nice for them to have a positive project too, something to do after all the dragons have been slain.

Cautionary optimists have proposed one possible positive project for experimental philosophers, namely to identify appropriate stimuli and controls so that intuition-pumping can lead to substantive philosophical conclusions after all. Some have labeled this “the positive program” of experimental philosophy.<sup>15</sup> As will become clear, I think there are other—better—positive programs for experimental philosophers besides nit-picking about intuitions, so I don't want to label this “the” positive program.

One potential concern is that Cautionary Optimists have yet to overcome the objections of Nay-Sayers. Cautionary Optimists are aware of this challenge, but are (perhaps unduly)

---

<sup>14</sup> As this paper has surely made clear, I don't trust intuitions. However, I have no official views regarding just how surprising and variable our intuitions are, so I can't predict how many more successes Nay-Saying experimentalists will end up enjoying. Nay-saying experimental methods can detect *inconsistency* which is one strong indicator of *unreliability*, but, as I argued in section 2, we should also doubt the reliability even of consistently held intuitions. Any successes that Nay-Sayers enjoy will provide all the more reason not to trust intuitions, but my own mistrust of intuitions doesn't depend upon nay-saying experimental results.

<sup>15</sup> Papers that draw a similar distinction between positive and negative programs include Alexander et al (forthcoming), Alexander & Weinberg (2007), Kauppinen (2007), Nadelhoffer & Nahmias (2007), and Weinberg (2007).

optimistic that, with appropriate controls, we will eventually be able to see past all the noise and context-sensitivity in people's intuitions and detect a steady signal. Regardless, there is a second deeper concern which has received less attention. As I mentioned in the preceding section, most current theories in philosophy of mind and philosophy of language hold that reference is determined, in large part, by external causal factors, and we have no reason to expect our intuitions to be strongly correlated with these factors. If this is right, then *even if* Cautionary Optimists manage to find just the right stimuli and just the right controls to detect a steady signal in people's intuitions, there is no reason to think this signal will be *correct*. For example, cautionary experimental philosophers might find that people have a very robust intuition that water is a homogenous substance, but that wouldn't show us that water *actually is* a homogenous substance. Similarly, even if cautionary experimental philosophers find that, after we set up certain controls, people have a very robust intuition that free action is incompatible with determinism, that won't show us that free action *actually is* incompatible with determinism. To presume that our intuitions would be good guides to these matters is to presume a theory of reference, like descriptivism, which holds that reference is determined by internal factors that our intuitions are likely to track reliably. But such theories of reference are no longer widely held.

So, I would encourage Cautionary Optimists to think carefully and critically about their (perhaps unknowing) commitment to a broadly descriptive theory of reference, given the many challenges to descriptivism from philosophers of mind and language and from Nay-Sayers, and especially given the plausible alternatives that are available. Even if Cautionary Optimists end up deciding to stick with a descriptivist package, they will likely benefit from examining and articulating their reasons for approaching things in this way.

In the meantime, I would encourage experimental philosophers of all stripes to consider pursuing alternative positive projects that would fit better with contemporary theories of reference. If experimental philosophers were to explore the causal or pragmatic factors which—unlike intuitions—actually do determine the reference of our philosophical concepts, then experimental findings could help us to draw positive philosophical conclusions about the referents of these concepts. Two such positive approaches are especially attractive:

*Naturalized Experimental Philosophy* allies itself with causal/informational theories of reference and with ‘naturalized’ approaches to philosophical problems.

*Pragmatic Experimental Philosophy* instead allies itself with teleo/pragmatic theories of reference and with Pragmatic Conceptual Analysis.

Both of these ‘package deals’ hold that the correct analyses of our concepts are determined in part by empirical facts about the causes and/or effects of our concept-usage. Hence, both can benefit from having clever experimenters help reveal what exactly these empirical facts are. I will first consider how experimental philosophers could advance a Naturalized research program, and then how they could advance a Pragmatic one.

#### 4. ‘Naturalized’ Experimental Philosophy

I can see at least two plausible ways that an experimentalist might contribute to a causal/naturalized approach:

**Step 1.** An experimentalist might help to determine what things play (or have played) an appropriate role in causing our use of a concept.

**Step 2.** An experimentalist might help to identify the relevant natural kind(s) that those things belong to.

These contributions might require experiments quite different from the intuition-pumping that experimental philosophers traditionally have done. For example, if, roughly following Gareth Evans (1973), one holds that a concept refers to the natural kind whose members have been the

dominant source of the information people associate with that concept, then, in Step 1, experimentalists would need to identify what objects have actually been the dominant source of that information. These would probably be objects that we're intuitively willing to apply the concept to<sup>16</sup> so some intuition-pumping might be helpful. However, (a) our intuitions about the far-out cases philosophers often consider (like trolley problems, True Temp, or worlds where supercomputers can predict all actions) will be quite irrelevant. And, (b) experimentalists will need to look not just at our intuitions, but also at the world, to see which, out of all the things we would be intuitively willing to apply a given concept to, are the things that actually have been sources for information we associated with the concept.

Of course, more work would need to be done to figure out exactly how experimentalists might contribute to this Evans-inspired approach. And of course there are other (perhaps more plausible) versions of a causal/naturalized approach besides this simple Evans-inspired one, and these other versions might call experimenters to focus upon other aspects of what has been causing our use of concepts or words. I won't dwell on these details, because I want to move on to the teleo/pragmatic "package deal" that I favor. But I've at least offered fans of causal/naturalized approaches a glimpse of a potential positive future for Experimental Philosophy, a future much brighter than the nay-saying and cautionary nit-picking about intuitions that has characterized so much Experimental Philosophy to date.

## **5. A Pragmatic Experimental approach to Free Will**

To illustrate how experimental philosophers might contribute to a teleo/pragmatic philosophical methodology, I will consider two examples in some detail: free will and (in the

---

<sup>16</sup> Unless we've regularly applied a concept in ways that, upon reflection, we find counter-intuitive—a real possibility.

next section) explanation. Let's begin with an intuition about free will reported by Frank Jackson, who, you'll recall, is a staunch advocate of Intuitive Conceptual Analysis:

Speaking for my part, my pre-analytic conception of free action is one that clashes with determinism. (Jackson 1998, p. 44).

Nay-saying experimental philosophers have pointed out one potential flaw in basing a theory upon Jackson's incompatibilist intuitions: many ordinary subjects do not share these intuitions (Nahmias et al, 2005, 2006). Other more cautionary experimental philosophers have held that when ordinary subjects depart from Jackson's intuitions, they are driven to do so by emotional factors, which some people take to be grounds for dismissing these dissenting intuitions (Nichols & Knobe 2007). This debate then gets mired in questions about which intuitions we should trust: do emotions *cloud* our intuitive judgment, or is emotion an essential part of our *appropriate* response to some cases?<sup>17</sup> While experimental philosophers might shed very interesting light on descriptive/psychological questions about how subjects arrive at the intuitive responses they do, there's little hope of directly connecting this descriptive story to normative/philosophical questions about what we can *correctly* count as a free action. To make philosophical progress here, experimental philosophers will need to take a different tack.

Let us focus, instead, upon a second potential flaw in basing our theory of free will on intuitions, a flaw that Jackson notices himself. Explications that satisfy our intuitions may fail to be useful. Worse, they may even force us to abandon the useful work that we have been using a concept to do. For example, given that determinism is true (or near enough to true) Jackson's incompatibilist explication of free action will say that *no* actions are free. This will undercut many useful distinctions that we ordinarily use the concept of free action to make, like the

---

<sup>17</sup> For much discussion of these and related questions, see Nichols & Knobe (2007) and Nahmias & Murray (forthcoming).

distinction between the “free” acts of ordinary vandals and the “unfree” acts of sleep-walkers.

Faced with this failing, Jackson proposes a fallback strategy:

It is, thus, only sensible to seek a different but ‘nearby’ conception that does, or does near enough, the job we give [to the concept being analyzed] in governing what we care about, our personal relations, our social institutions of reward and punishment, and the like, and which is realized in the world. (Jackson 1998, pg 45)

Jackson’s fallback strategy is a version of the methodology I call Pragmatic Conceptual

Analysis. It’s convenient when a key spokesman for Intuitive Conceptual Analysis admits that

it’s “only sensible” to use my proposed methodology instead of his! Perhaps experimental

philosophers could be “only sensible” as well, and focus their efforts upon revealing what

beneficial roles our concepts play for us, so that we may seek explications which will enable our

concepts to do this work even better.

Experimentalists might make at least two contributions to this pragmatic methodology, roughly analogous to the two listed for a Naturalized methodology above:

**Step 1.** An experimentalist might help to determine how we behave differently, depending upon whether or not we’ve applied a particular concept to something.

**Step 2.** An experimentalist might help to identify the ways in which these behavioral differences have regularly yielded beneficial outcomes.

In the case of free will, some existing experimental work might be construed as a crude initial attempt at Step 1. Experimental subjects who are given authoritative assertions that all people completely lack free will are then more likely than controls to be aggressive rather than helpful (Baumeister et al, 2009), to cheat on exams, and to steal money (Vohs & Schooler, 2008). This suggests that attributions of free will to oneself and others may play some role in generating pro-social behavior. However, it’s implausible that very many people ever in their ordinary lives make such blanket denials of free will (to *every* action of *every* person!). Instead, most ordinary attributions or denials of free will involve individual cases (“Did Chris freely choose to break my

sculpture?") or to certain classes of people ("Does methadone addiction prevent free action?").

To discover the useful work that attributions of free will serve in our everyday lives, we'll need a more nuanced approach.

To do this, experimentalists may design stimuli that will manipulate whether or not subjects will apply their concept of free action to some particular action, and then look for behavioral differences that depend upon this manipulation. One plausible hypothesis is that subjects might be more likely to advocate punishment for people they think act freely, than they do for people they think act unfreely. To test this hypothesis, experimentalists might give subjects something like the following prompt:

Randal was caught stealing CD's from a local music store. Randal has a history of shoplifting, and has been diagnosed as a kleptomaniac. After a thorough interview, an independent psychiatrist determined that Randal [freely chose / did not freely choose] to steal these CD's. After hearing all these details, a judge ordered Randal to spend a month in [prison / a mental hospital]. On the following scale, indicate how strongly you agree or disagree with the judge's decision:

1	2	3	4	5	6	7
Strongly Disagree						Strongly Agree

Our hypothesis would predict that subjects will agree more with a prison sentence when the shoplifter is said to have acted *freely* than when he is said to have acted *unfreely*, while conversely subjects will agree more with his being sent to medical treatment when he is said to have acted *unfreely* than when he is said to have acted *freely*. I'm fairly confident that this particular prediction will be borne out, but this is an empirical question so I could be wrong. (I hope someone will try it and see!) Even if the results in this case are quite foreseeable, in many cases, for many concepts, it will not be easy to foresee from the armchair how people's behavior hinges upon their application of concepts, so experimentalists may play a vital role in discovering these patterns.

This experiment pumped quite different intuitions from those pumped by traditional experimental philosophers: not intuitions about when to apply our concept of free action, but instead intuitions about what to do *after* we've applied that concept. Ultimately though, Pragmatic Conceptual Analysis is interested in how people *actually do* behave when they apply a concept, and intuitions about behavior are relevant only insofar as they reflect actual behavior. Especially for topics that are complicated or emotionally gripping, one might worry that intuitive responses to vignettes won't reflect people's behavior in real-life cases. To circumvent this worry, pragmatic experimental philosophers could run "ecologically valid" experiments in which subjects' actual behavior is witnessed. For example, subjects might be placed in the role of jurors listening to the details of Randal's case, and experimenters could observe whether jurors actually would recommend different sentences depending upon whether or not they are told a defendant acted freely. So, while pragmatic experimental philosophy might reasonably start by pumping (a special sort of) intuitions, it could easily lead to behavioral experiments strikingly different from those of traditional experimental philosophy.

These differences are heightened when experimentalists move on from (Step 1) discovering how people behave differently depending upon whether or not they've applied a particular concept, to (Step 2) discovering how these behavioral differences regularly turn out to be beneficial. I have no general account of "beneficialness" to offer,<sup>18</sup> but I trust that, at least in some cases, it is uncontroversial what is beneficial. For example, in the present case, I take it to be uncontroversial that it would be beneficial to achieve (at minimal cost) an outcome in which

---

<sup>18</sup> For many purposes, we'll want to construe whatever we think we have reason to pursue as "beneficial". If we construe "benefits" in this way then Pragmatic Conceptual Analysis will yield explications that will help us more consistently to achieve what we think we have reason to pursue. For certain explanatory purposes, it may be useful instead to construe as "beneficial" whatever is rewarded and/or selected for by the learning and/or evolutionary processes that have shaped the conceptual systems in question. If we construe "benefits" in this way, then Pragmatic Conceptual Analysis yields explications that enable us to give good explanations for the characteristic patterns of behavior and success and failure that these conceptual systems produce. See Fisher (2006, in prep Ch. 3).

people like Randal no longer engage in shoplifting. Here researchers can set up experiments to determine which people would most beneficially be punished, and which people would most beneficially be given medical treatment. (Notice that these experiments would be far different from those of traditional experimental philosophy.<sup>19</sup>) If it turns out that harsh prison sentences do well to deter people like Randal from shoplifting, then Pragmatic Conceptual Analysis will likely end up counting Randal's shoplifting as a free action. If it instead turns out that kleptomaniacs like Randal aren't really deterred by the threat of imprisonment but do respond well to medical treatment, then Pragmatic Conceptual Analysis will likely end up counting Randal's shoplifting as unfree.

It is fairly plausible that many of our actions stem from normal deliberative processes that are sensitive to expected punishments and rewards, while other actions (including, perhaps, Randal's) stem from other processes that aren't so sensitive to punishments and rewards, but are susceptible to treatment. If the experimental results do indeed turn out this way, then Pragmatic Conceptual Analysis will yield a compatibilist explication that counts the former actions as free and the latter as unfree—perhaps something like “*x* is a free action if and only if *x* is produced by normal deliberative processes”. Once again though, it is an empirical question whether or not our categorizing actions as free and unfree really does yield benefits in this way—it's an empirical question whether we use attributions of freedom to decide between punishment and treatment, and it is an empirical question when (if ever) deterrence and/or treatment really work. Rather than attempting to judge these issues from the armchair, it will be good to have philosophically

---

<sup>19</sup> Indeed, one might hope to find relevant research on these topics from the various non-philosophers who have studied deterrence. However, in many cases, philosophers will find that the philosophically relevant experimental questions won't have yet been asked. E.g., much of the research on deterrence has asked which *people* are generally deterrable. In contrast, folk usage allows that a given person may produce some actions that are free and others that are unfree. Understanding the usefulness of this practice will require asking which *actions* are deterrable in a given (sort of) person. When we need data that aren't yet available from other experimenters, philosophers may need to get involved in designing experiments. (We'll return to such division-of-labor questions in the concluding section.)

sophisticated experimentalists do the hard empirical work that it will take to see which proposed philosophical analyses are best capable of yielding the sorts of benefits our use of concepts like FREE ACTION have yielded.<sup>20</sup>

Above, I raised two sorts of criticisms against Intuitive Conceptual Analysis. One was that Intuitive Conceptual Analysis sometimes yields explications other than those we would have most reason to adopt. (It was this consideration that led Jackson to concede that it is “only sensible” to employ something like my pragmatic methodology instead.) In contrast, if experimental philosophers can help us to uncover which proposed explication of free action would best sustain the ways that our counting some actions as “free” has been beneficial, then we will have instrumental practical reason to adopt this explication to continue achieving benefits in those ways. So my proposed pragmatic version of experimental philosophy comes out ahead of cautionary intuition-pumping on this count.

Second, I argued that there is little reason to expect intuition-pumping to yield explications that *correctly* characterize the actual referents of our concepts, both because our intuitions (especially about extraordinary cases) have a shoddy track record, and because our most widely accepted (causal/informational and teleo/pragmatic) theories of reference hold that reference is determined in a way that is quite independent of our intuitions. It is outside the scope of this paper to argue that a teleo/pragmatic theory of reference is right. However, *if* such a theory is right (as I and many others<sup>21</sup> believe is the case), then empirical findings about regular beneficial usage would be highly relevant to determining how we can correctly characterize the actual

---

<sup>20</sup> It may be worth noting that this enquiry could yield some surprising results. For example, suppose it turns out that offering rewards and punishments does affect the activities of certain sleep-walkers (or phobics, or obsessive-compulsives, or hypnotized people, or drug addicts) in much the same way that it does ordinary “free” actions. If so, then Pragmatic Conceptual Analysis would likely end up counting these activities as “free”, regardless of intuitions to the contrary. Similarly, it might turn out that, surprisingly enough, many of our habitual (or culturally inculcated, or even heroic) activities aren’t susceptible to rewards and punishments in the right ways for them to count as “free”, regardless of intuitions to the contrary.

<sup>21</sup> See note 5.

referents of our concepts. E.g., if a teleo/pragmatic theory is right in holding that “free action” refers to whatever natural kind has been playing an appropriate role in sustaining our regular beneficial usage of this concept, then empirical findings about regular beneficial usage can help us correctly to characterize the nature of free action. Hence, at least insofar as we have reason to think a teleo/pragmatic theory of reference is more likely to turn out to be right than is the sort of old-fashioned descriptivist theory that would support intuition-pumping, my proposed pragmatic experimental methodology comes out ahead on this count too.

## **6. A Pragmatic Experimental approach to Explanation**

Let us now move on to consider a second example: explanation. This example is especially appropriate for two reasons. One reason is that one of the clearest sustained attempts to employ a version of Pragmatic Conceptual Analysis is Jim Woodward’s (2003) book on causal explanation. Woodward writes:

[M]y project focuses on (what I take to be) the purposes or goals behind our practices involving causal and explanatory claims. [...] [W]e introduce concepts (including concepts of cause and explanation) and characterize them in certain ways at least in part because we want to *do* things with them [...]. [Proposed explications of these concepts] can be well or badly designed for such purposes and we can evaluate them accordingly. (Woodward 2003, pp. 7-8)

Woodward thinks the primary purpose of our concept of a good explanation is to identify explanations that “furnish information that is potentially relevant to manipulation and control: they tell us how, if we were able to change the value of one or more variables, we could change the value of other variables” (Woodward 2003, p. 6). Woodward then fashions a detailed manipulationist account of explanation and thinks we should embrace it because it does especially well at fulfilling this purpose.

One might wonder whether or not Woodward correctly identified, from his armchair, the purposes served by our concept of a good explanation. This brings us to the second reason why

explanation is especially appropriate for our purposes. For many experimentalists have taken up the question of what purposes explanations serve for us. Their work may serve as a model for how experimental philosophers can make positive philosophical contributions, not by pumping intuitions about fanciful cases, but instead by exploring the useful things concepts actually do for us.

You may notice that Woodward and I both shifted freely between the usefulness of our concept of a good explanation and the usefulness of explanations themselves. There is typically a clear distinction between the usefulness of a concept and the usefulness of that concept's referent. Man-eating tigers might be useful for various things, but those things are quite distinct from what *a concept of man-eating tigers* is useful for: avoiding getting eaten. This distinction breaks down when a concept is used primarily to identify items we will then employ for further uses, for a full account of the usefulness of such a concept will contain little more than an account of the usefulness of its referent. Such a breakdown occurs in the case of good explanation. The primary use of our concept of a good explanation is to identify which proffered explanations to retain and use as explanations.<sup>22</sup> So, to learn the purposes served by this concept, we'll need ask what purposes are served by the things we count as good explanations.

There has been a great deal of experimental work directed at this question (much of which is summarized in Lombrozo 2006). I'll briefly describe four strands of this research, to give a flavor for how it might proceed.

First, there is evidence that knowing of a good explanation that links two properties together makes subjects more likely to attend to potential statistical relations between those properties. Subjects will attend to statistical relations that they can explain in a given dataset (e.g., a correlation between cars' sizes and their mileage), but will overlook or ignore statistical relations

---

<sup>22</sup> This, itself, is an empirical claim, and could be subject to experimental testing.

they can't explain (e.g., correlations between color and mileage). But when subjects are antecedently given an explanation for how two properties might be linked (e.g., a car's color might inspire the driver to accelerate more aggressively, yielding worse mileage) they then notice and attend to the statistical relations between those properties (Koslowski 1996). This suggests that good explanations do not just enable prediction and manipulation, but also guide our epistemic activities as we decide what information to gather and attend to.<sup>23</sup>

Second, explanations play a role in how subjects will generalize information to novel cases. If subjects can't explain why members of a particular class (e.g., Kehoe ants) display a particular trait (e.g., a stinging bite), then they will attribute that trait to things outside that class based on their overall similarity to members of that class. But if subjects *can* explain why members of the class display that trait (e.g., in terms of levels of iron sulfates in the ants' blood), then subjects will attribute the trait to things outside that class, not on the basis of *overall* similarity, but instead on the basis of whether or not those other things share whatever features explained that trait in the given class (Rehder 2006). This suggests that we use good explanations not just to understand the causal operations of particular systems, but also to help generalize our existing knowledge to novel systems.<sup>24</sup>

Third, what explanations subjects have in mind affects how they will categorize various items. In an experiment by Tania Lombrozo (2009), subjects are told that, in a particular category of things (e.g., a particular species of flower), feature C (e.g., a chemical in the flower's stem) often causes another feature E (e.g., the flower bends over). Half the subjects are prompted to explain E *functionally* (e.g., bending over helps the flower spread its pollen) while others are prompted to explain E *mechanistically* (in terms of its cause, C). Subjects were then asked to rate

---

<sup>23</sup> Additional evidence for this comes from experiments on illusory correlations, e.g., Wright & Murphy (1984).

<sup>24</sup> For further evidence, see Sloman (1994).

the probability that things with just C or just E (but not the other) were members of the category in question. Subjects who explained E mechanistically are more likely to take the presence or absence of the cause C to be diagnostic of whether a thing belongs to the category, whereas subjects who explain E in terms of the function it serves are more likely to take the presence or absence of the functional effect E to be diagnostic of membership in that category.<sup>25</sup> This is evidence that counting something as a good explanation may significantly affect later categorization, so a full understanding of the purposes that explanations serve will probably need to take this usage into account.<sup>26</sup>

And, fourth, there is a great deal of evidence that having, and especially generating, good explanations helps to facilitate learning. Subjects who are encouraged to generate explanations are better at remembering both presented facts (Alevan and Koedinger, 2002) and procedures (Calin-Jageman and Ratner 2005) than are subjects who engage in alternative study methods. Furthermore, as you might expect after Rehder's experiments above, there is strong evidence that explanations help to facilitate generalization of learned material to applications involving novel cases.<sup>27</sup> This work suggests that part of the point of distinguishing good explanations from bad is that good explanations present information in a way that is well-suited to learning and generalization.

These four strands of experimental research suggest that Woodward's estimation of the purposes served by explanations is incomplete. I don't doubt that Woodward is right that

---

<sup>25</sup> This may involve some sort of default presumption that functions are multiply realizable, so that category-members who don't realize the functional effect E the most common way (via C) should instead realize it in some alternative way; whereas mechanistic explanations take an effect to be comparatively unlikely to occur in a category member without its normal cause.

<sup>26</sup> See also Ahn (1998), Jameson & Gentner (2008).

<sup>27</sup> This has been evidenced regarding learning in multiple domains, including physics (Chi et al., 1989), biology (Chi et al., 1994), mathematics (Wong et al, 2002; Rittle-Johnson 2006), folk psychology (Amsterlaw & Wellman, 2006), and even fictional robots (Williams & Lombrozo, forthcoming).

(1) explanations provide us with information that aids in prediction and manipulation.<sup>28</sup>

However, these other strands of research suggest that explanations also (2) highlight statistical regularities worth attending to, (3) help us to generalize information to novel cases, (4) help us to categorize various items, and (5) facilitate learning and memory. For an explanation to do all these jobs well, it will need not just to provide the sort of rich causal information that Woodward describes, but also to package this information in a way that appropriately fits with patterns of human curiosity, generalization, categorization, and learning.

This suggests two conclusions about explanation. First, notice that Woodward's identified purpose for explanation will tend to count explanations as better the more detailed causal information they provide. In contrast, purposes involving the details of human psychology will tend to count explanations as better if they are relatively brief and manageable, fore-fronting the information that is most relevant to us. Our psychology deals better with forests than catalogs of trees. Elsewhere (Fisher 2006, in prep, Ch. 5) I have argued that good explanations are able to fulfill these conflicting purposes (to be both richly informative and manageably brief) by effectively having the structure of a hyper-linked document like a web-page. A good explanation prominently highlights a simple and highly relevant pattern of dependence, but it also offers hyperlinks (or footnotes, or psychological associations, or key words) that will lead to further information about why this pattern of dependence generally should hold, about how this dependence may be practically exploited, and about the details of the particular explanandum in question.

Second, notice that Woodward's identified purpose (providing recipes for manipulation and intervention) will tend to count an explanation as better the more closely it fits the details of the particular system being explained. In contrast, the other purposes identified above include ways

---

<sup>28</sup> Some defense and empirical support for this hypothesis is given by Gopnik (1998).

in which explanations lead us to generalize to *other* systems, or to categorize borderline cases (like ones that display an effect without its normal cause) which don't have the same detailed causal structure as the system being explained. This suggests that we should count an explanation as better if it highlights a pattern of dependence that holds, not just in the given system, but also will generalize nicely to related systems. Elsewhere both Tania Lombrozo (2009) and I (Fisher 2006, Ch 5) have argued that this suggests that there is something importantly right about unificationist understandings of explanation.

One might worry whether it will be possible to explicate 'good explanation' in a way that will enable this concept to fulfill all the different purposes it has been serving for us. This is a general challenge, closely akin to the challenge facing fans of traditional Intuitive Conceptual Analysis when they encounter conflicting intuitions. There are several ways of meeting this general challenge, including bifurcating a concept, splitting the difference, abandoning some prior uses, or relativizing the concept to particular uses. (These general options and their merits are discussed in Fisher (2006, in prep, Ch. 3.) In the case of explanation, I have argued (Fisher 2006, in prep, Ch. 5) that relativization is appropriate. So, for example, what counts as a good explanation *relative to the needs and interests of the local Weight-Watchers society* might be different than what counts as a good explanation *relative to the needs and interests of a conference on mammalian metabolism*. So, I've proposed, our account of good explanation will say that, *other things equal*, an explanation is better if it is more informative, if it is more correct in its details, if it affords more practical interventions, if it is more easily manageable by human psychology, and/or if it generalizes smoothly to a broad range of cases. But we often won't be able to optimize all these dimensions simultaneously, so we'll often settle for explanations that are very good relative to some dimensions but not others.

I don't pretend that this work is completed. There are many open questions about exactly what purposes our concept of good explanation serves, and about how exactly we should explicate this concept to best serve those purposes. So there is much more work for both experimentalists and more traditional philosophers to do. But the current research already indicates that we'll need an account that allows that good explanations have some sort of hyper-linked structure, and that allows for different tradeoffs between rich causal information, manageability, and generalizability. Hence, even while there remains work to do, good progress has already been made, and this progress may serve as a positive model for how experimental work that explores the actual usefulness of concepts (rather than pumping intuitions about their applicability) might contribute to our understanding of a philosophical concept.

## 7. Conclusions

I have considered three different package deals combining views in Philosophy of Mind and Language, Meta-Philosophy, and Experimental Philosophy. One of these packages gave center-stage to pumping intuitions about fanciful cases, but that package involved problematic commitments both to a controversial descriptivist theory of reference and to following intuitions that nay-saying experimental philosophers have shown to be suspiciously variable and context-dependent. In light of these difficulties, it would be good for future-minded experimental philosophers to align themselves with a different package deal—a positive project to pursue once all the dragons are slain. I have suggested two such options. Experimentalists might help fans of “Naturalized” approaches discover what natural kinds have been playing an appropriate role in causing us to use concepts as we do. Or—preferably, in my eyes—experimentalists might instead help practitioners of Pragmatic Conceptual Analysis discover how our concept usage regularly yields beneficial outcomes, so that we can then craft philosophical analyses that will

enable us to yield such beneficial outcomes more consistently. Analyses that preserve and enhance the useful work that our concepts have been doing will be analyses worth adopting, and finding such analyses is a worthy positive project for all philosophers (both experimental and traditional) to adopt.

I've argued that a great deal of experimental research may be highly relevant to philosophers' attempts to explicate philosophical concepts. One might wonder whether such experimental research is "real philosophy", or whether it should instead be counted as psychology or sociology or some other science. One might wonder why we philosophers should muddle around trying to do experiments ourselves, when we could instead do what Jesse Prinz (2008) calls *empirically informed philosophy*, and just draw upon the results of experimentation by real scientists.

Empirically informed philosophy is great, as far as it goes. In the case of *explanation*, we were lucky to have a great deal of philosophically relevant evidence from non-philosophers, largely because of the co-incidence between the usefulness of *our concept of explanation* (which is of special relevance to philosophers) and the usefulness of explanations *themselves* (which was of more direct interest to psychologists). However, this was a lucky co-incidence. For most philosophical concepts, we won't have such a co-incidence to draw upon.

Furthermore, when we're trying to analyze notoriously thorny philosophical concepts (like freedom, knowledge, explanation, justice, or meaning), the relevant empirical facts about our usage will likely be very difficult to tease out, especially in a way that can arbitrate between the subtle differences in proposed philosophical analyses. *Explanation* is one of the best cases for empirically informed philosophy, but, as I noted above, even in this case the current experimental results aren't enough to fully decide which proposed account of 'good explanation' is best. In the case of *free will*, we may hope to tap into existing research on the efficacy of

various treatments and deterrents, but it's much less likely that we'll find existing research showing us exactly how our use of this concept is connected to the various deterrents and treatments that have been studied.

Things are even worse for other philosophical concepts. Consider the many subtly different analyses of *knowledge* proposed in response to the Gettier problem, including analyses that differ only in minor details involving multiple layers of defeaters and defeater-defeaters. It would be great if psychologists would spontaneously do enough experiments on the usefulness of our concept of knowledge to reveal which of these proposed analyses would best sustain our patterns of beneficial usage. But, we shouldn't hold our breath. Scientists are unlikely to spontaneously generate all the results that philosophers will need as soon as we would want them.

For this reason, it is important that philosophers take an active role in designing experiments geared towards delivering the subtle results that we'll need to decide between competing philosophical analyses. We might do this on our own or (often better) in collaboration with experienced scientists. Either way, we need experiment-designers who understand philosophy as well as experimentation. In other words, we need *experimental philosophers*, not just armchair philosophers (even empirically informed ones) and social psychologists.

The future of experimental philosophy can be positive and bright, if experimental philosophers will focus on the right empirical questions. The examples described in this paper—free will and explanation—can serve as good models for a worthwhile positive program:

Pragmatic Experimental Philosophy.

## ACKNOWLEDGEMENTS

I am grateful to helpful comments from many people, including Shaun Nichols, Ron Mallon, Joshua Knobe and other participants at the 2009 NEH summer institute on experimental

philosophy and the 2011 bootcamp on experimental philosophy of free will. I am also grateful for helpful feedback from my colleagues at SMU, and audiences at the Eastern APA (including helpful commentary from Thomas Nadelhoffer), the Southern Society for Philosophy and Psychology, and Florida International University.

## REFERENCES

- Ahn, W. (1998). "Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality." *Cognition* 69: 135–78.
- Aleven, V. and Koedinger, K. (2002) "An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor." *Cognitive Science* 26: 147–79.
- Alexander, J., and J. Weinberg. (2007) "Analytic epistemology and experimental philosophy." *Philosophy Compass* 2: 56–80.
- Alexander, J., R. Mallon, and J. Weinberg. (forthcoming) "Accentuate the Negative." *Review of Philosophy and Psychology*.
- Amsterlaw, J. and Wellman, H. (2006) "Theories of mind in transition: a microgenetic study of the development of false belief understanding." *Journal of Cognitive Development* 7: 139–72.
- Appiah, A. (1986) "Truth Conditions: A Causal Theory." In *Language, Mind and Logic*, Thyssen Seminar Volume, Jeremy Butterfield (ed.) Cambridge University Press, pp. 25–45.
- Baumeister, R.F., E.J. Masicampo, and C.N. DeWall. (2009) "Prosocial benefits of feeling free: Disbelief in free will increases aggression and reduces helpfulness." *Personality and Social Psychology Bulletin* 35: 260–268.
- Blackburn, S. (2005) "Success Semantics". In *Ramsey's Legacy*, H. Lillehammer and D. H. Mellor, eds. Oxford University Press.
- Boyd, R. (1988) "How to be a Moral Realist," in *Essays on Moral Realism*, ed. Sayre McCord. Cambridge University Press, pp. 181–228.
- Burge, Tyler. (1979) "Individualism and the Mental." in P. French, T. Uehling, and H. Wettstein, eds., *Studies in Metaphysics*. University of Minnesota Press.
- Calin-Jageman, R. and Ratner, H. (2005) "The role of encoding in the self-explanation effect." *Cognition & Instruction* 23: 523–43.
- Chalmers, D. (2002) "The Components of Content" in David J. Chalmers ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford UP, pp. 608–633.

- \_\_\_\_\_. (2004) "The Foundations of Two-Dimensional Semantics" in M. Garcia-Caprintero and J. Macia, eds. *Two-Dimensional Semantics: Foundations and Applications*. Oxford University Press.
- Chi, M., M. Bassok, M. Lewis, P. Reimann, and R. Glaser. (1989) "Self-explanations: how students study and use examples in learning to solve problems". *Cognitive Science* 13: 145–82.
- Chi, M., N. de Leeuw, M. Chiu, and C. Lavancher. (1994) "Eliciting self-explanations improves understanding." *Cognitive Science* 18: 439–77.
- Craig, E. (1990) *Knowledge and the State of Nature*. Oxford University Press.
- Dretske, F. (1988) *Explaining Behavior*. MIT Press.
- Evans, G. (1973) "The Causal Theory of Names." *Proceedings of the Aristotelian Society*, Suppl. vol 47: 187-208. Reprinted in *The Philosophy of Language*. Ed. A.P.Martinich. Oxford University Press, 1996.
- Fisher, J. (2006) *Pragmatic Conceptual Analysis*. University of Arizona dissertation.
- \_\_\_\_\_. (forthcoming). "Meanings and Methodologies." To appear in *New Waves in Philosophy of Mind*. J. Kallestrup and M. Sprevak, eds. Palgrave MacMillan.
- \_\_\_\_\_. (in prep). *Pragmatic Conceptual Analysis*. Near-completed book manuscript.
- Fodor, J. (1990) *A Theory of Content and Other Essays*. MIT/Bradford.
- Frege, G. (1892) "On Sense and Nominatum" in A. P. Martinich, ed., *The Philosophy of Language*. 3rd ed. Oxford: Oxford University Press, 1996. pp. 186-98.
- Gopnik, A. (1998) "Explanation as orgasm." *Minds and Machines* 8: 101–18.
- Griffiths, P. (1997) *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Griffiths, P. and K. Stotz. (2008) "Experimental Philosophy of Science." *Philosophy Compass* 3: 507-21.
- Haslanger, S. (1999) "What Knowledge Is and What It Ought To Be: Feminist Values and Normative Epistemology." in *Philosophical Perspectives* 13: 459-480.
- Jackson, F. (1998) *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. New York: Oxford University Press.
- James, W. (1906) "What Pragmatism Means."  
<http://www.marxists.org/reference/subject/philosophy/works/us/james.htm>

- Jameson, J., & Gentner, D. (2008). "Causal status and explanatory goodness in categorization." In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *In Proceedings of the 30th annual conference of the cognitive science society*. 291–6.
- Kauppinen, A. (2007) "The rise and fall of experimental philosophy." *Philosophical Explorations* 10: 95–118.
- Kornblith, H. (2002) *Knowledge and Its Place in Nature*. Oxford University Press.
- Koslowski, B. (1996) *Theory and Evidence: The Development of Scientific Reasoning*. MIT Press.
- Kripke, S. (1972) *Naming and Necessity*. Cambridge: Harvard UP.
- Lewis, D. (1972) "Psychophysical and theoretical identifications." *Australasian Journal of Philosophy* 50: 249–58.
- \_\_\_\_\_. (1984) "Putnam's Paradox." *Australasian Journal of Philosophy* 62: 221-236.
- Lombrozo, T. (2006). "The structure and function of explanations." *Trends in Cognitive Sciences* 10: 464–70 .
- Lombrozo, T. (2009) "Explanation and categorization: how 'why?' informs 'what?'" *Cognition*, 110: 248–53.
- Machery, E., R. Mallon, S. Nichols, and S. Stich. (2004) "Semantics, Cross-cultural style." *Cognition* 92: B1–B12.
- Mallon, R., E. Machery, S. Nichols, and S. Stich. (2009) "Against arguments from reference." *Philosophy and Phenomenological Research* 79: 332–56.
- Margolis, Eric and Stephen Laurence eds. (1999) *Concepts: Core Readings*. MIT Press.
- Millikan, R. (1984) *Language, Thought, and Other Biological Categories*. MIT Press.
- \_\_\_\_\_. (2000) *On Clear and Confused Ideas*. Cambridge, UP.
- Nadelhoffer, T., and E. Nahmias. (2007) "The past and future of experimental philosophy." *Philosophical Explorations* 10: 123–149.
- Nahmias, E., S. Morris, T. Nadelhoffer, and J. Turner. (2005) "Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility." *Philosophical Psychology* 18: 561-84.
- \_\_\_\_\_. (2006). "Is incompatibilism intuitive?" *Philosophy and Phenomenological Research* 73: 28–53.
- Nahmias, E., and D. Murray. (forthcoming) "Experimental philosophy on free will: An error theory for incompatibilist intuitions." in *New Waves in Philosophy of Action*. Palgrave.

- Nichols, S. and J. Knobe. (2007) "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions." *Nous* 41: 663-85.
- \_\_\_\_\_. (2008) "An Experimental Philosophy Manifesto." In J. Knobe and S. Nichols (ed.), *Experimental Philosophy*. Oxford UP.
- Papineau, D. (1987) *Reality and Representation*. Blackwell.
- Prinz, J. "Empirical Philosophy and Experimental Philosophy," in J. Knobe and S. Nichols (Eds.), *Experimental Philosophy*, Oxford University Press (2008).
- Putnam, H. (1973) "Meaning and Reference." *The Philosophy of Language*. Ed. A.P.Martinich. New York: Oxford UP, 1996.
- Rehder, B. (2006) "When similarity and causality compete in category-based property generalization." *Memory and Cognition*. 34: 3–16.
- Rittle-Johnson, B. (2006) "Promoting transfer: the effects of direct instruction and self-explanation." *Child Development* 77: 1–15.
- Roskies, A. and S. Nichols. (2008) "Bringing Moral Responsibility Down to Earth." *Journal of Philosophy* 105/7: 371–88.
- Russell, B. (1905) "On Denoting." in A. P. Martinich, ed., *The Philosophy of Language*. 3rd ed. Oxford: Oxford University Press, 1996. pp. 199-207.
- Ramsey, F. (1927) "Facts and Propositions." In *The Foundations of Mathematics, and Other Logical Essays*. R.B. Braithwaite, ed. London: Routledge and Kegan Paul, 1931, pp. 138-55.
- Rupert, R. (1999) "The Best Test Theory of Extension: First Principle(s)". *Mind & Language* 14: 321-55
- Slovan, S. A. (1994) "When explanations compete: The role of explanatory coherence on judgments of likelihood." *Cognition* 52: 1-21.
- Strawson, P.F. (1950) "On Referring." in A. P. Martinich, ed., *The Philosophy of Language*. 3rd ed. Oxford: Oxford University Press, 1996. pp. 215-30.
- Swain, S., J. Alexander, and J. Weinberg. (2008) "The instability of philosophical intuitions: running hot and cold on truetemp." *Philosophy and Phenomenological Research* 76: 138–55.
- Vohs, K.D. and J.W. Schooler. (2008) "The value of believing in free will: Encouraging a belief in determinism increases cheating." *Psychological Science* 19: 49–54.
- Weinberg, J., S. Nichols, and S. Stich. (2001) "Normativity and epistemic intuitions." *Philosophical Topics* 29: 429–60.

- Weinberg, J. (2007) “How to challenge intuitions empirically without risking skepticism.” *Midwest Studies in Philosophy* 31: 318–343.
- Whyte, J. (1990) “Success semantics.” *Analysis* 50: 149–57.
- Williams, J.J. & Lombrozo, T. (forthcoming) “The role of explanation in discovery and generalization: evidence from category learning.” *Cognitive Science*.
- Wong, R.M.F. et al. (2002) “The effects of self-explanation training on students’ problem solving in high-school mathematics.” *Learning & Instruction* 12, 233–262
- Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.
- Wright, J. and Murphy, G. (1984). “The utility of theories in intuitive statistics: The robustness of theory-based judgments.” *Journal of Experimental Psychology: General* 113: 301–24.