80-818 Keith Douglas
w/ Horacio Arló-Costa kdouglas@andrew

Normative, Descriptive, Idealized: Metaphilosophical Reflections on Epistemology

Introduction

Traditionally, epistemology is regarded as a normative rather than a descriptive discipline[1]. Nevertheless, papers and monographs in epistemology often criticize the works of other epistemologists for being too unrealistic in one respect or other. The goal of the present paper is to articulate what ways exist in order to evaluate an epistemological work for relevance to actual agents[2].

This paper will consist of three parts. The first part will consist of a discussion of how the "extreme views" do not do the work we want and move into a more general discussion. A second part will discuss the difference between idealizations and normativity. A third section will attempt to locate some resources to answer the questions we have raised. Since this paper is only intended to raise a problem, we shall be content with only this discussion, leaving a more definitive solution of it for another time. Nevertheless we briefly canvass some possible approaches to solutions in section three.

Broadly speaking, our project is a first blush at one of consilience sensu Whewell or Wilson (1998). We wish to know how to integrate epistemology with the other fields that study reasoning, belief, and so forth.

[1] We shall remain agnostic on whether there is legitimately a descriptive portion to epistemology (or indeed, a normative one). This paper will hopefully clarify the problems with claiming too much for either side of this (to be argued, false) dichotomy.

[2] There are some epistemologists who claim that their work is also applicable to other things besides humans. For instance Gärdenfors (1988) claims the possible applicability of his work to databases. Churchland (2002) discusses the possibility of her work being applicable to other animals. We shall remain agnostic on this point throughout, though we will have the opportunity to discuss points at which the category the agent falls under might make an approximate difference.

Before we begin, however, we must forestall one major source of disagreement. It has occurred to the present author that many of the facets of the debates discussed in this paper might very well have questions related to the debates over realism as overtones. We do not wish to enter this issue at the present time and will leave it be with a single remark. We have nothing to say to someone who would claim a massively antirealist or non-realist position. That is, we take it for granted that there is some attempt, even in normative or idealized epistemology to "get things right." (The realist can substitute "true" for "right" in the above vague expression, but should not hinge her answers in the debates below on this notion.)

Part 1: Locating the Problem

The problem we are concerned with can be raised as follows. We are interested in demarcating psychology (or "cognitive science" broadly construed) from epistemology, and also to see what sorts of resources are available to argue for or against adopting certain theses in epistemology. Recent programs such as "Bayesianism" are attractive, but raise questions of idealization and so forth that are difficult to tackle without a well developed metaphilosophy.

We note that issues similar to our concerns have been raised by Weinberg, Nichols and Stich[3]. They suggest that there is a grave problem in traditional approaches to normativity in epistemology. They put this as follows (italics in original):

> "Reflective equilibrium strategies and other Intuition Driven Romantic strategies all yield as outputs claims that putatively have normative force. These outputs tell us how people ought to go about forming and revising their beliefs, which belief forming strategies yield genuinely justified beliefs, which beliefs are warranted, which count as real knowledge rather than mere opinion, etc. But there is a problem lurking here – we'll call it the *Normativity Problem: What* reason is there to think that the output of one or another of these Intuition Driven Romantic strategies has real (as opposed to putative) normative force? Why should we care about the normative pronouncements produced by these strategies? Why should we try to do what these outputs claim we ought to do in matters epistemic? Why, in short, should we take any of this stuff seriously? "

[3] We note that one does not need to adopt the extreme epistemic relativism promoted by Stich to agree with his conclusion in the present passage.

The important part of Weinberg, Nichols and Stich's remarks for us is their "normativity problem". Why should we adopt a particular normative pronouncement, considering there are many competing claims to the "correct" answer? Weinberg, Nichols and Stich go on to show that psychological investigation demonstrate the possibility is not merely conceivable, but one that actually obtains. (We do not endorse all of the arguments these authors provide against their critics: some are off topic. For example, the possibility that the subjects' judgments are not reflective enough or result from enough discussion can be dismissed against the subjects of this paper. This reply would simply be *ad hominem* against professional philosophers and other researchers, and is such not pertinent.)

Another somewhat related concern has been raised recently by Patricia Churchland (2002). She has pointed out the difference between what she calls the Platonic strand in epistemology as opposed to the Aristotelian. The former is characterized by speculation and *a priori* investigation; the latter by actual investigation of what cognizes, perceives, etc. These positions are best understood as extremes and (as she realizes) do not completely accurately describe the historical Plato and Aristotle. Note also that this division does not necessarily divide into normative/descriptive in quite our fashion. Nevertheless her emphasis that one should be careful about *a priori* investigation is apt. (Of course, we do not assume that normative programs in epistemology are necessarily *a priori* in any particular sense the reader may have in mind. We do not wish to investigate the thorny history of that term.)

It is also important to note that we do not intend to argue there should be a "Berlin wall" of separation between epistemology and psychology. Instead we are concerned that appeals to idealizations or to normativity may hide unrealistic epistemological programs. Or, worse still, give us no principled means to evaluate claims (for or against) of unrealism. This latter point applies to both descriptive attempts **and** attempts to provide normative accounts. A descriptive account does presuppose some normative ideals. We turn

to some cases to this effect now.

Sometimes it is said that it is relatively straightforward to (e.g.) demarcate psychology from epistemology because the former does not deal with normative questions. This is not quite right, as Haack (2000 [1978]) pointed out in the context of logic. (As we shall see below there is a parallel between our question about epistemology and questions about logic.) Psychology does make normative judgments, for instance about illusions of perception. (It might be argued that these considerations presuppose some epistemological views, e.g. a denial of radical skepticism or that they presuppose a weak realism. We grant this, however, this just makes the problem of locating epistemology in the "intellectual map" more acute.) This extends even to social psychology, where theories of cognitive dissonance require some notion of conflicting belief. Conflict here is usually taken in some usually ill-specified logico-semantic sense.

Furthermore, there are fields strongly related to psychology that are also normative, but in a different sense. Such fields might include psychiatry (both technology and practice[4]), clinical psychology (again, both technology and practice) education, marketing, management, etc. It seems pretheoretically (unless we wish to play "Humpty Dumpty" with the term "epistemology") that these disciplines are also not what we have in mind by our subject. Epistemology (to the extent that it is) is normative in a different sense. Any answer to our problem must account for this,

[4] Following Bunge (1983) we recognize the existence of the psycho- and socio-technologies listed. Further, we also classify the design of medical procedures as a biotechnology. This is important as it reinforces the plausibility of the existence of psychotechnologies provided one believes there is an important biological component to psychology. See also Douglas 2002 for more on psychotechnologies from a slightly different perspective.

even if only to ultimately dismiss the distinction[5].

It is probably safe to say that if we value epistemology at all we will not be content with pure description[6]. This should be rather obvious, given that we cannot realistically describe without some sort of framework (or better, theory) to describe what we are observing. On the other hand, the other extreme, doing epistemology of a perfectly omniscient knower is also an unrealistic move.

But what is considered an "adequate" position has changed over the years. Plato in (e.g.) the Meno at 85b-d stresses the importance of the individual and goes so far as to claim that knowledge is all "internal" in a very radical (to the contemporary reader, at least) sense:

> "Socrates: What do you say of him, Meno? Were not all these answers given out of his own head?
>
> Meno: Yes, they were all his own.
>
> Socrates: And yet, as we were just now saying, he did not know?
>
> Meno: True.
>
> Socrates: But he still had in him those notions - had he not?
>
> Meno: Yes.

---

[5] There is a practice (which we shall not discuss further) called "philosophical counseling" which claims to replace or supplement traditional psychiatry and the like. It is the view of the present author that this practice amounts to quackery, but this is a story for another time. However, it is important nevertheless to point out that this field would further muddy the normative-descriptive waters in so far as it purports to provide "world view" counseling and other **modifications** of behaviour and belief.

[6] Even Patricia Churchland (2002; see above) is interested in normative questions in epistemology. She, however, feels it is very important to get the details of the descriptive questions as correct as possible before normative questions are put on the table. Her book is (in part) an attempt to show what sorts of details might be necessary from neuroscience in that respect.

> Socrates: Then he who does not know may still have true notions of which he does not know.
>
> Meno: He has.
>
> Socrates: And at present these notions have been stirred up in him, as in a dream; but if he were frequently asked the same questions, in different forms, he would know as well as anyone, at last?
>
> Meno: I dare say.
>
> Socrates: Without anyone teaching him he will recover his knowledge for himself, if he is only asked questions?
>
> Meno: Yes.
>
> Socrates: And this spontaneous recovery of knowledge in him is recollection?"

Plato has Socrates say that coming to learn is recollection, that done within individual in himself with a minimal of outside influence. He even says that the slave boy will come to **know** (and for Plato that is a very strong condition indeed!) with a minimal amount of outside influence.

Descartes' approach, at least in the *Meditations* and similar writings was analogous[7]. This emphasis on the solitary inquirer (despite his large correspondence!) was criticized implicitly *avant la lettre* by Bacon (in the *New Organon* and the *New Atlantis*, with their calls for scientific societies) and yet continued by both him and those who followed such as Locke, Hume, Berkeley and

[7] It is not clear to the present writer if the more scientific writings of Descartes (e.g. *Le Monde*, *Dioptrique, La Géometrie* etc.) are as individualistic in their epistemology as the epistemology espoused in the more traditionally philosophical works. For lack of time in the current project, the author will remain agnostic over this point.

Kant[8]. Thus this view of epistemology was shared by those labeled by history as rationalists and by those called empiricists alike.

However, by the late 20th century, we (finally?) arrive at texts such as Goldman's *Knowledge in a Social World* (1999) which stress the community of knowers as playing decisive roles[9]. It is now (more or less) accepted that the epistemic community is of vital philosophical consideration. How much so is of course subject to some debate, but all the matters for the present purpose is the change in the idealization. Interestingly enough, this may have arisen as a result of a change in epistemic norms **within** knowledge communities - i.e. the increase in the necessity of collaborative research in science, etc. as problems become more and more challenging. This historical point would be an interesting way to further extend this research, but not one that the author has any time to explore in enough depth at the present time.

While we have illustrated our thesis by means of the social context of knowledge, it is applicable to any epistemological concern. We shall adopt the family of views called "Bayesianism" as a point of departure for some of what follows, as it can illustrate a great deal of issues surrounding idealization, normativity and other concerns, as we shall see.

---

[8] The Royal Society (and perhaps other scientific societies) was inspired by Bacon's social proposals. Nevertheless, the individuals commonly regarded as philosophers in the (approximately) Baconian tradition did not adopt this viewpoint. Increasingly, however, figures more on the scientific camp (Boyle, Hooke, etc.) felt this was a good idea and thus formed the Society. (Of course, it must be said that not all the scientists of the period even in England recognized its usefulness: Newton's support seems to have been grudging.)

[9] It is sometimes stated that the fashionable view (in some circles) of stressing the social network as the **only** thing of epistemological relevance (e.g. by Bruno Latour and the so-called "Strong Programme" of Barnes and Bloor) is the equally extreme counter reaction to the extreme individualist epistemology of previous centuries. There is insufficient information (and inclination at this time on the part of the present author) to investigate this at this time, but it would nevertheless prove interesting.

For purposes of this paper, (following Arló-Costa's report on Seidenfeld, private communication) we shall take "Bayesianism" to be the family of views which assert that:

(B1) An agent's belief state is represented by a coherent, finitely additive conditional probability[10] P[K] (.|.)

(B2) P[K] (.|.) is relativized to background evidence K (consistent and closed under entailment).

(B3) As regulated by Bayes' theorem for conditional probability P[K] (.|. & e) is the agent's hypothetical belief state for the hypothesis that he accepts only the new consistent evidence e, i.e. under the hypothesis that K is enlarged by addition of e (and its consequences given K).

As is well known, there are numerous places in which this remarkably broad spectrum of viewpoints is criticized. One way is that it if "subjective probabilities" are elicited from most agents, they fail to meet requirement (a), as has been repeatedly noted in the literature. (See, for example, Osherson's [1998 {1995}] review article.)

Partisans of the Bayesian approach have retreated from the descriptive claim and claim that although it appears humans are very far from being Bayesian agents, they nevertheless ought to be in some sense. Here we have a typical move of the kind we are concerned with. Is this move acceptable?

Part 2: Normativity and Idealizations

As noted above, the Bayesian rescues his hypotheses by claiming they are normative. In particular, he tells us that the Bayesian program can even show that the individual "ought" to tailor the strengths of his beliefs in the "probabilistic" way the view suggests. He presents the Dutch book argument (or something

[10] This characterization does hide that a large number of very different understandings of probability - even from purely the formal perspective - are at work here in different Bayesian programs.

similar) to suggest that our rationality (as determined pretheoretically[11]) is at stake, and so forth.

But is this a hypothesis of normativity or one of idealization? Both seem plausible, and yet both potentially run into problems.

One problem with the claim of normativity concerns idealization itself. The move from normativity to idealization is illustrated next. In the sense that an omniscient being would know a superset of the set of all human knowledge, human knowledge could be idealized in some contexts to be assumed to be perfect. If this strikes the reader as a ridiculous assumption to make, she should recall that in some contexts it is made in select domains. For instance, typical presentations of "belief dynamics" (e.g. Gärdenfors 1988) presuppose logical omniscience on the part of the agents described. Gärdenfors justifies this postulate in the following way (pp. 22):

> "It is clear, however, that the two criteria are not realistic as descriptions of individuals actual sets of belief. In particular, this is the case for the requirement of including logical consequences - because of the limitations of our mental powers, we often do not see all the consequences of what we accept. I still believe, however, that the criterion is useful, at least as an ideal of rationality."

Is this so? Gärdenfors' system also abstracts away from the time variable. If that is to be considered important in an epistemological model, it is not immediately clear that being logically omniscient is always an ideal of rationality because of the so-called "frame problem." (See Dennett 1998 [1984][12].) We shall return to these and other considerations later in our "laundry list" below.

---

[11] *Nota bene!* This condition is a great example of the subject of this paper. The Dutch Book argument presupposes that being a "money pump" is irrational, or that one who rejects certain kinds of bets is irrational. (I.e. the argument presupposes that the agent will bet on anything whatsoever.)

[12] The gist of the argument here is that one should not be logically omniscient as one would spend one's time deriving irrelevant consequences from one's beliefs and never act.

If the Bayesian wants to postulate a parallel ideal of "probabilistic omniscience", she also runs into similar issues. For instance, she must confront whether the idea of computing posterior probabilities via (say) Bayes theorem, or Jeffrey conditioning etc. is realistic. After all, humans do not **explicitly** calculate these sorts of things very often. Of course, there have been claims about different forms of computation going on without our awareness of such in various works from the usual forms of "computational theory of mind" to Marr's (1982) extreme (but interesting) claim that the visual system "computes" a convolution and a gradient.

In short, by claiming that her postulates are normative, the Bayesian runs quickly into the (admittedly vague) Kantian principle familiar to students of ethics: "ought implies can." In other words if it is (grossly) unrealistic[13] to expect an agent to be capable of doing something, he surely cannot be faulted for not trying. Hence, we must be very careful when claiming normativity to save a system of postulates.

Thus if the claim is instead that the Bayesian postulates are an idealization, rather than normative in character, she runs into an analogous series of problems. In particular, she must ask what exactly is being idealized. For instance, the usual Dutch Book argument is synchronic; diachronic Dutch books (which thus involve the time variable and are more realistic in some respects) are racked with conceptual difficulties even for those who are "more Bayesian" than the present author. (See, for example, Levi 2002 [2000].)

Focusing on just the Bayesian approach can yield a suitable laundry list of assumptions that in principle could be challenged by the enterprising critic. We list these to illustrate the

[13] What "can" is supposed to mean in the principle is the primary source of vagueness. In particular, it is not clear how much an agent has to be "capable" of doing to count as fulfilling this requirement. For example, both in the ethics case and the epistemological one it runs afoul of the free will vs. determinism debate, and in particular such issues as the "could have done otherwise" maxim.

diversity of possible points of disagreement, and to show how the question we are debating is very much an issue.

The laundry list is not meant to be a list of things that epistemologists as a **whole** have forgotten. The questions are rather: which of these features are necessary to have an appropriate epistemological theory? Which of these are legitimate idealizations (for any given use) - say for normative considerations?

The first of these is simply the psychological observation that the structure of credences does not match the axioms of the probability calculus, as we previously noted. This point needs further elaboration. We make three interrelated remarks:

(a) We acknowledge the existence of (for example) what Kahneman and Tversky (1979) have called "prospect theory" which attempts to describe the behaviour of decision making under risk and the 'deformation' to probability functions that results. Clearly we do not mean to imply that Bayesians are unaware that the axioms of the the probability calculus are being used as an idealization (or perhaps as a normative consideration).

(b) The above said, however, it is important to note that the critical question with idealization still remains, even with prospect theory on the table. We ask again: why do we think that the Bayesian approach (i) is sufficiently "close" to what real agents can do that it is a good idealization, and (ii) hence why do we think it might have normative force.

(c) Hence also how do we evaluate debates over these issues? It is true that Kahneman and Tversky (for example) recognize the **difference** between idealized (and hence possibly normative) and descriptive. But they do not address the question of whether the idealized is "close enough." A decision theoretic example may make this clearer. Suppose[14] that someone invented a decision theory that would give correct answers to **any** decision problem (i.e. a

[14] The author is indebted to brief discussions with his colleague, Jiji Zhang, for developing this example.

completely utopian notion), but was such that no matter how it was used (by computer or people or whatever) it could only yield its answer after the choice was in fact made. It seems reasonable that this decision theory would be regarded as normatively correct (*ex hypothesi*) but otiose or useless.

A second is the assumption (used even if Bayesianism is taken normatively or as an idealization) that credence strength suitably gets mapped linearly to the interval [0,1]. Another assumption is that credence strength is a continuous quantity (or at least, is sufficiently "close" to one so as to not produce large errors.) Another concerns the lack of the time variable in the theory of probability proper. (To many, this raises immediate suspicion that it would have nothing to do with anything factual without further ado.) We also are not told what credences are (or "subjective probabilities") or what their **change** is. A behaviourist would no doubt want them to be some form of disposition to behave as is her usual wont; a more neuropsychologically inclined individual would suggest they are certain types of brain processes; a functionalist that they are a species of functional organization, and so forth.

We see with this example that epistemological hypotheses are not even isolated within epistemology. The last consideration above quickly runs into metaphysical concerns. The stuff of which the credences are properties (etc.) can be of vital importance if one is theorizing as to their changes, even phenomenalistically. (That is, without regards to mechanism: the ideal gas law familiar from freshman science classes is a phenomenalistic equation.)

In addition to this, there is a computational question: each time the agent needs to update "her" probability function, she must perform a **calculation** unconsciously. This seems to commit the Bayesian to a weak form of computationalism in the philosophy of mind.

Another collection of considerations (of the ones dreamt up on the spot by the present author) concerns the hypothesis that each agent possesses only one probability function. Conceivably, each agent could have two or more domains with each their own

probability functions. This model would explain why some people have different epistemic standards in different domains (e.g. scientific and religious beliefs). One might also wonder about developmental concerns: when does a human being acquire her probability function[15]? How? Other concerns that one might have pertain to the domain over which the probability functions are to be applied: propositions, sentences, events[16], etc. If the domain of the probability function is taken to be something sentence-related (i.e. either sentences or propositions) then the question of whether non-sentence-like items (emotions, percepts, etc.) can have any bearing on it also arises.

Before turning to approaches that might solve our problem, we must examine briefly another suggestion on our theme. Arló-Costa (personal communication) has suggested that:

> "Epistemology deals centrally that deals with the theory that permits the agents to 'see' that their judgments were incorrect."

We think this is right, but doesn't quite solve our problem, because it still does not specify the "difficulty" of the standard. Again, logical omniscience would be nice, and most people who are taught logic recognize that logical omniscience is thereby a standard to attempt to live up to[17] . But just because one recognizes that in one case a certain norm is the correct one and that one cannot live up to it completely, it does not follow that any norm we cannot live up to and recognize our errors in terms of is legitimate without further ado.

---

[15] Piaget, for instance, raised the question of the relationships between the adult knower and the child.

[16] Here "event" is taken in its metaphysical sense, not as the generic term for whatever is in the domain of the probability functions. See Douglas 2001 for more on this potential confusion in the literature.

[17] In certain contexts! The frame problem (discussed above) makes living up to the logically omniscient ideal problematic. We teach this constrained aspect to logic to students as well when we tell them that not all inferences are worth drawing in a given context.

Arló-Costa is right to suggest we need normative background theory to tell us **when** agents make errors. But what sort shall we choose? How many errors are they "allowed to make" and of what character? Many people criticize other proposals in epistemology because they do not meet certain broadly speaking "empirical" considerations. For example, Haack (1993) discusses the infinite regress argument that is supposed to provide the impetus to foundationalist approaches in epistemology. She points out (implicitly) that infinite epistemic regresses are only a problem for finite agents. So, is it an **error** that one commits when one does not have epistemic resources to finish solving a problem?

Of course humans are commonly thought to be finite (including by the present author!). But that does not entail (without further ado at any rate) either that (a) modeling us as infinite is undesirable or that (b) those who thought us to be infinite[18] are (again, without further ado) completely mistaken. Haack (1993, pp. 107) also discusses whether we (humans) have beliefs about colour patches in our visual field. Some people would regard this specific consideration about our beliefs to be irrelevant to epistemology.

Some might argue that there is no real debate about the above considerations[19], so let us take an example where there is much disagreement about specifics. Consider reflecting on the nature of choice - why should we think that "mere" reflection gets us to the right place? This is the Stich *et al* worry. Well, of course if near univocality can be obtained, that would work. But this obviously doesn't occur, with no end of debates in sight, we think

[18] Gödel, famously, though that it might be possible that the human mental states could diverge to infinity over its (infinite) existence. (See, e.g., Wang 1974, 1987.) Gödel believed in life after death, so this infinity is not a problem on **his** grounds; but here again we run into a metaphysical question infecting epistemological considerations.

[19] It should be noted that the fact that one could disagree in **principle** on the lines given should be enough to demonstrate the point. We nevertheless move to other examples to illustrate the variation of problems we are concerned with.

this proposal is a bit of a pipe dream. It might be argued that the basic principles are agreed upon and details are what is debated. The problems with this claim are: (a) how many people have to agree that we've (for now) captured the right notion - there **are** dissenters to "the" theory of rational choice, for instance; (b) what counts as basic principles in order for us to **tell** when agreement (of the appropriate degree) is reached? In other words, we have merely pushed the problem back a step. An improvement, but we are not out of the woods yet under this proposal.

As for people coming **to see** that they made an error, this is an interesting approach. Of course, it is a bit vague. How does one tell this? Obviously if a subject in an experiment spontaneously volunteers that she sees that she was in error, and that she failed to live up to some normative principle or other, that's a pretty decent indicator of the appropriate sort of seeing. But what about other cases? Do we do a "forced choice" experiment after a belief experiment to determine whether or not people felt they should have "seen" something? This has some merit - it would tell us at least what people **feel** they ought to be capable of. This of course is rather different from what they ought. It does not even tell us that the principle is reasonable, though. Consider the parallel case of a religious ethical system that postulates that all humans are depraved sinners and yet nevertheless should **aspire** to be morally perfect. Many religious believers such feel this way. (I.e. that the striving is what is regarded as important, etc.) To those those who do not share such religious beliefs such an approach is sometimes regarded as dehumanizing and in itself immoral because (for example) of the cognitive dissonance it provokes.

Note also that there are differing reports about the degree to which people recognize their (putative) errors. Dawes (1988, pp. 36) reports that people stand by their choices in experiments designed to elicit framing effects. On the other hand, Raiffa (1961), did see that he would change his mind in face of the evidence that he was incoherent. (Or rather, he said he would act differently in a situation involving real money, which is

something rather different.) Raiffa had studied the Allais paradox prior to his experience, which might account for the difference here. Perhaps the two different sorts of studies are sufficiently different that people react different to normative considerations in both cases. However, the fact that there are **some** cases where agents do not "grasp the normative force" is sufficient for **some** doubt.

Of course, these reports and other considerations do not directly consider epistemological modeling but instead theories of rational choice. However, due to the fact that they involve (for instance) psychological considerations relating to probability judgments, they thus infest our theme in that way. The aforementioned consideration is especially important for the Bayesian, for obvious reasons. (We have taken the Bayesian family of epistemologies as a test case for some remarks, but our aim is broader.)

Part 3: Debating Solutions and a Tentative Suggestion

We have focused on epistemology previously, but the concerns of the present paper could be extended to other fields one might think have a normative vs. descriptive vs. idealization "issue" as well. There are at least two others in this category: logic and ethics. Ethics will take us too far afield, and too far outside the present author's expertise. However, the comparison with logic is apt, and does have a reasonably substantive literature. This section will start with this comparison then debate some other solutions to our problem.

We begin with Haack's remarks. As noted above in section one, Haack (2000 [1978]) has tried to explore the normative/descriptive dichotomy in the context of logic. We have suggested previously that both normative and descriptive aspects to epistemology also require another ingredient, idealization. Like epistemology, logic also has idealizations in both its normative and descriptive aspects.

For instance, taken as a description of how we ought to reason, the first order predicate calculus is adequate in the sense that

we teach it with that in mind to undergraduates around the world in hundreds of colleges and universities. But even here[20] we run into issues we have discussed previously, in particular "computationalism", "propositions", etc. In fact, Haack points out that the usual view that logic is about propositions seems to run into problems of subjectivism. Propositions are understood to be "private" entities, sentences more public, so she suggests instead using sentences as the domain of logical laws produces a certain way in which logic can be rendered more objective[21].

But does this help us understand our concerns in epistemology? Simply saying that there are analogous problems in logic does not help. It does propose one solution, however. That is, it suggests, along the lines of the teaching logic to undergraduates, that the problems are apparent, not actual. We manage to teach logic, so the problems (if there are any) are not genuine problems. This of course is just a "sweep it under the rug" solution, and is of course exacerbated by the fact that there is less consensus on what constitutes elementary epistemology. Another possibility that it suggests is moving more of the epistemological concerns "out in public view." It is not clear how one would go about doing this without presupposing some sort of weak logical behaviourism (that beliefs are just "manifest on the surface"), and, as might be surmised, that would be severely question begging.

Another possibility, also suggested by logic, would be to list all of the presuppositions of one's epistemological account[22]. This is unsatisfactory for several reasons. First, it presupposes that one can do so in a reasonable amount of time and space: the laundry list assembled above shows how an apparently simple principle can have many presuppositions. Second, it presupposes we know exactly

[20] I.e. even ignoring the complications raised by modal logics, paraconsistent logics, higher order logics, etc.

[21] Of course, this creates other problems that we do not need to go into in the present paper.

[22] For instance, one could axiomatize it. This does raise regress problems of the sort that axiomatics in logic produces, but we shall not deal with this here.

what the presuppositions are, which is far from true in many cases. Third, even if we could list all the presuppositions, it would not solve the problem. It would simply give the critics of the "unrealism" a place to "look up" the points at which they disagree, not give them a principled way in order to do so. A more "axiomatic" approach to epistemology would suffer from the same flaws, regardless of its other substantial benefits.

Let us now turn away from logic for suggestions and look at other possibilities. Another possible answer would be simply to drop either the normative or the descriptive end of epistemology[23]. No one has seriously proposed the latter[24]; we need not consider it for that reason. However, the move to drop the normative is (on one interpretation at least) the goal of the school of "naturalized" epistemology which stems from the work of Quine[25].

The attraction here would be that one **always** rules out projects that are normative; this way there would be no arbitrariness in the judgment of the realism in that respect of projects and so forth. However, it would suffer from three problems. The first problem would be that it simply makes epistemology itself otiose. The second objection would lie in that it is arbitrary to rule out all normative epistemology to begin with. The third, and best argument, is the argument presented by Kim (1993 [1986]).

This argument is worth examining in some detail as it speaks not only to the current consideration, but the considerations of our

[23] It is interesting to note that "logic" in English is slowly losing its descriptive meaning, whereas in French it is less strange to say phrases like "logique de la paix" and so forth.

[24] To be fair: Richard Rorty has arguably defended dropping all of epistemology, so a fortiori he has defended dropping descriptive epistemology as well. We will not address this here.

[25] Patricia Churchland (see above) considers herself in that tradition, but even she does not go so far as to repudiate normative epistemology altogether as we have noted previously. Even Quine himself is deeply ambiguous on this point. See Haack 1993 for a discussion.

present paper as a whole. First, Kim considers belief to be an action of sorts. This yields immediately the consequence that assuming we want normative ethics[26], we want normative epistemology, because normative epistemology is just a focused portion of normative ethics. Kim takes normative epistemology to be focused on justification, and attempts to show further that if justification is lost (as it would be in a purely descriptive epistemology), knowledge itself would disappear from epistemology's consideration. An epistemology without the concept of knowledge would be a rather strange epistemology indeed[27].

Kim explains that Quine's approach focuses on causal relations, as opposed to evidential ones. Since this radically changes the subject, Kim continues, it is not clear how one (traditional vs. naturalized epistemology) could be in competition with the other, never mind one be a replacement for the other.

We can extract another moral from Kim's argument. First is the interesting consideration of action. Kim is not the only one to link knowing and doing; the theme is a well known approach of the pragmatist schools. This suggests the approach that any proposal of (normative) epistemology must explain how **it** is to be **used**. This might appear to make epistemology more technological (see Douglas 2002) than might be previously thought. However, since the proposals for action are not for other ends but (perhaps) for the sake of knowing, we need not draw this conclusion. This has the pseudoPopperian merit that we can then see what would verify or falsify the epistemological program. We ask of the model questions like: "Does this allow the agent to minimize error?" or the like and investigate it as if it were a theory of science. This recovers some of Quine's proposals for naturalized epistemology,

[26] An assumption which the present paper will not even begin to attempt to defend. We nevertheless remark on the assumption (one not noted by Kim himself!) in order to play by our own rules.

[27] Though not unheard of; Pyrrhonian skepticism is arguably one such. Of course, this school has as a view that there is no knowledge, not that knowledge is not worth investigating. Richard Jeffrey (1968) is arguably one modern partisan of this viewpoint, so it is not merely of historical interest.

as testing theories in epistemology would require good models in psychology. Further, as desired by some (the present author; Bunge [1983]) this would suggest no clear dividing line between psychology and epistemology. This latter result might be regarded as a reductio or at least an unfortunate feature. We shall not pursue this matter here[28].

It should be noted that this proposal links up nicely with Dennett's (1998 [1988]) suggestion that artificial intelligence be regarded as both a branch of engineering and as a branch of philosophy. His suggestion to take (certain) attempts at AI to be attempts to validate epistemological[29] hypotheses dovetails as a special case of the above proposal.

Another proposal that might work to solve our problem to some degree is suggested by Kelly (1990 [1985]). He tells us (pp. 120) that we can build models of idealized agents who incorporate in them theories of their very limitations. For example, he takes standard computability theory (sensu Turing, for instance) to include such. Presumably the proofs of the unsolvability of the halting problem and other limitative results are what he has in mind. He puts it this way (pp. 121):

> "The moral is this: if you insist on advising the public to hitch its wagon to a star, you ought to provide a way to characterize which broken wagons came closer to the stellar rendezvous than others and to explain why coming closer is better than not trying at all."

(*Pace* Kelly, we do not take it as necessary that one **proves** results about the limitations of an approach, however.) Further, Kelly also echoes Dennett's remarks which link artificial

[28] Hopefully, this work shall continue in a sequel to the present paper where this objection will be answered.

[29] They also validate (or invalidate) metaphysical hypotheses concerning the nature of mental states, etc. as well, but that's another story for another time. That said, it is important to realize that metaphysical hypotheses will enter into the test of epistemological ones, and it may prove difficult to extricate the two contributions in certain cases. We shall postpone this consideration to the sequel of this paper.

intelligence with epistemology. In fact, Kelly goes so far as to suggest that AI itself is more creditable when interpreted as investigations in effective (his terminology) epistemology. Nevertheless, there may be some who find fault with this approach. We find it likely, however, that there would be a metaphysical character to these debates. Suppose someone were to build a model of an agent who could use a learning-theoretic algorithm with, say, a polynomial running time to demonstrate, *pace* some philosophical critics, that the learnability considerations of a given viewpoint was indeed feasible. This raises the metaphysical question of whether human learning mechanisms are sufficiently like the (so-called) von Neumann machine architecture that the running time calculations were applicable[30]. More on metaphysical considerations later.

Another approach would be to build from the bottom up: develop a model that was so minimal that everyone would agree that it describes the least powerful epistemic agent. This is the approach taken by Cherniak (1986). From there one could conceivably add refinements piecewise, debating each as they come to publication, debate and so forth.

This approach, while initially attractive, returns us to one of our earliest concerns in this paper: what constitutes an agent. Further, it raises the question of whether the power of epistemic agents even be ordered in a relevant respect. For instance, some philosophers would consider only mentally competent adult humans (i.e. not adults in a persistent vegetative state) to be agents in the appropriate sense; some would include children or even infants. Others would include various nonhuman animals: other primates; all other mammals; some might include birds. Other

[30] We do not mean to suggest that any particular epistemologist (or AI researcher) is guilty of this error. Nevertheless, it is important to remember that (at least) two assumptions are built into the argument our imaginary learning theorist is making. First, that "computationalism", is broadly speaking a correct way of understanding how humans learn. Second, that the assumptions that go into running time calculations are correct for the "human computer." These assumptions do vary depending on the architecture of the computer in question (Cormen, Leiserson and Rivest 1990).

chordates might be included or not; some biologists would say that cephalopods (squids, octopuses, etc.), are more epistemic agents than reptiles, amphibians or bony fish, so they should be included. And of course some philosophers would include some computers, or their programs, or robots, or ...

Merely trying to find an ordering principle to include all of these above, nevermind one that would allow everyone to agree on the minimum, is a tall order. If such were actually done, we would reconsider this issue. Until then, however, we regard this problem as a substantial obstacle to the above proposal.

Another monograph on our theme is Goldman's (1986) volume. He makes some useful remarks that we should attempt to integrate in our discussion. The first of these is on page 5. He argues for the importance of understanding for epistemology how actual cognitive agents do perceive, remember, reason and so forth. This centers around their evaluation first and foremost. This seems to be the usual normative consideration. But these quickly run into questions of **what** is being evaluated. In order to see if our beliefs are structured in a certain way, it must be possible in the sense that our beliefs are the sorts of things which could have such a structure. So epistemologists must pay attention to cognitive science in order to learn and analyze what sorts of things beliefs are. This seems right, except for the fact that the notion of "possibility" here is a bit vague. Assuming one had access to a well developed theory of (say) nomological possibility, this approach would do the job in this case. As before, epistemology thus has an interesting metaphysical dependence. Goldman admits as much when he says that this approach ties his inquiry to topics in the philosophy of mind.

Further, Goldman (pp. 24) tells us that normative standards must also tell us when they apply. This seems useful, and fits in with the Dennett-Kelly point we have raised earlier. Finally, Goldman's discussion of "winner-take-all" neural networks as an alternative to the usual Bayesian model of coherence is interesting, and is motivated by some of the same considerations that the present author (2001) has raised concerning the latter approach.

Related to the above proposals would be to tie our normative epistemology to the nomological state space of the appropriate system. This (like any model) would involve an idealization. However, by suitably tying to the nomologically possible states of an epistemic agent one would obtain normative principles that were appropriate to it. We recognize the following problems with this proposal that would have to be solved for it to be worth while. One is that it makes explicit metaphysical commitments as to the lawfulness of epistemic agents. This might prove problematic on some philosophies of mind (e.g. that of Davidson 1980 [1974]). Two, the relevant time and space bounds might prove contentious, but at least they would be explicit. For example, the Turing formalism (see Douglas forthcoming) can be seen as normative in a certain respect: it tells us the normative standards of effective computability. But it does this through the imposition of (for example) finite (although unbounded) time of operation. (We presuppose here that any statespace for an epistemic agent will be modeled at least implicitly) with time and space variables.) Three, it requires that one agree on an understanding of law, which is by no means an easy task. Four, it quite likely presuppose much work in philosophy of mind, the cognitive sciences and so on that is just not available, or would be highly contentious if used.

The conciliatory parts present in this approach to our problem make it interesting. Alas due to the reasons we have canvassed above it is problematic for the same reasons.

Another, final, approach to solving our problem would be simply to deny that one should find a solution to the problem. Instead, one would leave everything as it is, and hope that tacit principles would allow us to settle the matter eventually. We dislike this approach as it goes against the principle of conscilience that we are working with, but if none of the other proposals above are successful we will have to deal with this possibility.

Conclusion

We have located our problem, debated its relevance and proposed

some first attempts at solving it. Future work shall expand upon these solutions and develop them further. We note in passing that two areas of importance (metaphysics, historical consideration) require substantial development outside of epistemology proper and may very well require more attention to more metaphilosophical than this paper has touched upon.

Bibliography

Bunge, M. 1983. *Treatise on Basic Philosophy*, Vol. 6: *Understanding the World.* Dordrecht: Reidel.

Cherniak, Christopher. 1986. *Minimal Rationality*. Cambridge: MIT Press.

Churchland, Patricia. 2002. *Brain-wise: Essays in Neurophilosophy.* Cambridge: MIT Press.

Cormen, Thomas, Leiserson, Charles, and Rivest, Ronald. 1990. *Introduction to Algorithms*. Cambridge: MIT Press.

Davidson, Donald. 1980. (1974). "Psychology as Philosophy." Reprinted in *Essays on Actions and Events*. Oxford: Oxford University Press.

Dawes, Robyn. 1988. *Rational Choice in an Uncertain World*. Orlando: Harcourt Brace Jovanovich.

Dennett, Daniel. 1998. (1984). "Cognitive Wheels: The Frame Problem of AI." Reprinted in *Brainchildren: Essays On Designing Minds*. Cambridge: MIT Press.

Dennett, Daniel. 1998. (1988). "When Philosophers Encounter Artificial Intelligence." Reprinted in *Brainchildren: Essays On Designing Minds*. Cambridge: MIT Press.

Douglas, Keith. 2001. "A Survey and Review: What is Probability?" Unpublished paper for class 80-600, Carnegie Mellon University, Fall 2001 semester.

Douglas, Keith. 2002. "Art, Technology and the Role of Aesthetic Values in HCI" Unpublished paper for class 80-990, Carnegie Mellon University, Fall 2002 semester.

Douglas, Keith. Forthcoming. *Turing's Analysis of Computation: Historical Development and Philosophical Discussion*. MS Thesis for program Logic and Computation, Carnegie Mellon University.

Gärdenfors, Peter. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge: MIT Press.

Goldman, Alvin. 1986. *Epistemology and Cognition*. Cambridge: Harvard University Press.
Goldman, Alvin. 1999. *Knowledge in a Social World*. Oxford: Clarendon Press.
Haack, Susan. 1993. *Evidence and Inquiry: towards reconstruction in epistemology*. Oxford: Blackwell.
Haack, Susan. 2000 (1978). *Philosophy of Logics*. Cambridge: Cambridge University Press.
Jeffrey, Richard. 1968. "Probable Knowledge." In *The Problem of Inductive Logic.* Lakatos, Imre, ed. Netherlands: North-Holland.
Kahneman, Daniel and Tversky, Amos. 1979. "Prospect theory: An analysis of decision under risk." Econometrica 47, pp. 263-291.
Kelly, Kevin. 1990. (1985). Effective Epistemology, Psychology and Artificial Intelligence." Reprinted in Sieg, Wilfried, ed. *Acting and Reflecting: The Interdisciplinary Turn in Philosophy*. Dortrecht: Kluwer Academic Publishers.
Kim, Jaegwon. 1993. (1986). "What is Naturalized Epistemology?" Reprinted in *Supervenience and Mind*, Cambridge: Cambridge University Press.
Levi, Isaac. 2002 (2000). "Money Pumps and Diachronic Books." In Barrett, Jeffrey and Alexander, McKenzie, eds. *Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association*.
Marr, David. 1982. *Vision*. New York: W. H. Freeman.
Osherson, Daniel. 1998 [1995]. "Probability Judgement." In *An Invitation to Cognitive Science, vol 3: Thinking*, Osherson, Daniel, ed. Cambridge: MIT Press.
Raiffa, Howard. 1961. "Risk, Ambiguity, and the Savage Axioms: Comment." The Quarterly Journal of Economics: **75**(4):690-694.
Van Fraassen, Bas. 1995. "Fine-Grained Opinion, Probability and the Logic of Full Belief." Journal of Philosophical Logic: **24**: 349-377.
Wang, Hao. 1974. *From Mathematics to Philosophy*. London: Routledge & Kegan Paul.
Wang, Hao. 1987. *Reflections on Kurt Gödel.* Cambridge: MIT Press.
Weinberg, Jonathan, Nichols, Shaun and Stich, Stephen. 2001. "Normativity and Epistemic Intuitions." Web version at:

<http://ruccs.rutgers.edu/ArchiveFolder/Research%20Group/Publications/NEI/NEIPT.html> ; forthcoming in Philosophical Topics.

Wilson, Edward. 1998. *Consilience: the unity of knowledge*. New York: Knopf.