Accuracy, Inquiry, and Belief

This paper concerns inquiry and its norms. Inquiry, and norms of inquiry, are often thought of as the project of science and scientists. But all of us engage in inquiry throughout ordinary life, wondering about banal questions, taking banal steps toward uncovering their answers. I wonder whether I have a can of chipotles in adobo sauce; I poke around the pantry to learn the answer. I wonder what the cat will do if I put my finger in his mouth while he’s yawning; I perform a simple but dangerous experiment. Conversation involves sharing information, narrowing down the space of possibilities in our epistemic common ground. 1 Conversations are often governed by questions under discussion, 2 and aim at collectively reasoning toward their answers. In general, we investigate our world, gather evidence, narrow down the space of epistemic possibilities, aiming to narrow our probability distributions around the actual world. 3

A central question on the norms of inquiry—“zetetic norms” 4 —concerns how they relate to epistemic norms. Epistemic norms have been thought of as governing beliefs and credences, while the norms of inquiry extend to actions. But both seem to concern learning about the world.

A common view of epistemic norms is what I’ll call “alethic consequentialism”: the view that the epistemic value of beliefs (credences, etc.) is a function of their accuracy—the more accurate, the better—and that rational beliefs are those that best conduce to that value. This view is often simply called “epistemic consequentialism”, because the view that accuracy is the source of epistemic value is so widely accepted. 5 (Accuracy-first epistemology is a popular recent version of this view. 6 ) And since the goal of inquiry is to learn truths, it’s natural to think of zetetic norms as alethic consequentialist norms as well.

---

2 See Roberts’s (1996) for an influential defense.
3 For a detailed discussion of the mental states and activities involved in ordinary inquiry, see Friedman (forthcoming).
4 I borrow this term from Friedman (2020).
5 There are some exceptions: Littlejohn (2015), for example, argues that knowledge, rather than truth, is the source of epistemic value.
6 Seminal works in this area include Joyce (1998, 2009); Greaves & Wallace (2006); Leitgeb & Pettigrew (2010a,b); Pettigrew (2016).
This paper argues against alethic consequentialism for both epistemic and zetetic norms. The general claim is this: we can come to have true beliefs in two distinct ways: by discovering truths and by creating truths. Only the former is epistemically or zetetically valuable. In §1, I give neutral characterizations of epistemic and zetetic norms and alethic consequentialism. In §2.1, I present a case study to motivate the rejection of alethic consequentialism in both the epistemic and zetetic cases. §2.2 considers various ways the example can be modified to avoid objections. In §3.1, I diagnose the problem for alethic consequentialism— that it conflates discovering and creating truth as sources of value—and elucidate the distinction. §3.2 discusses the relation between this diagnosis and a prima facie plausible theory about the relation between epistemic and zetetic norms. §4 zeroes in on the zetetic realm. While §3 argues that creating truth isn’t a source of positive zetetic value, §4 explores the hypothesis that it may be a source of positive zetetic disvalue. In §§5, I consider two possible explanations for why creating truths may be zetetically problematic: that it’s incompatible with conducting inquiry, and that it blocks off access to sources of genuine zetetic value. Both explanations face some potential concerns, but none amounts to a decisive objection.

1 Norms and consequentialisms: epistemic and zetetic

1.1 Epistemic and zetetic norms

“Epistemic norms” and “zetetic norms”, on my use, don’t just pick out whatever norms govern beliefs or inquiry, respectively. Epistemic and zetetic norms are specific flavors of norm with their own specific forms of goodness and rightness, distinct from prudential or moral norms (though not necessarily distinct from each other). Zetetic and nonzetetic norms apply to inquiry; epistemic and non-epistemic norms apply to belief. Often it’s (prudentially, morally, perhaps all-things-considered) best to do things that are epistemically or zetetically nonideal. From a zetetic perspective, for example, it’s best if you do nonstop research. But of course, it’s perfectly reasonable to rewatch a familiar movie or stare into space or take various other actions that do next to nothing to help you discover truths.

---

7 Under “norms”, I include various normative categories, including “oughts”, values, rational statuses, and so on.
8 For discussion see Friedman (2020); Thorstad (2021)
It’s conventional wisdom that there are distinctly epistemic norms.\(^9\) So far, no conventional wisdom in the younger literature about zetetic norms has emerged. One might hold that the norms governing inquiry are the same as the norms governing any other behavior: moral or all-things-considered norms.\(^{10}\) Why think there are distinctively zetetic norms?

First, actual scientific inquiry offers many cases of perceived conflict between moral and zetetic norms, cases where it feels as though we’re pulled in competing directions. There can be better and worse acts of inquiry that are morally on a par: for example, better and worse experimental designs with no morally relevant differences. Moreover, there can be cases where some action is zetetically good but morally bad: for example, experiments where live octopodes are subjected to starvation or shocks by electric prods. Such experiments are unlikely to afford significant downstream benefits that morally justify them. There’s been pushback from scientists against animal welfare laws, e.g., the European Union’s Directive 2010/63/EU on the “Protection of Animals used for Scientific Purposes”, which extends to cephalopods the legal protections earlier granted only to vertebrates. Some scientists argue these laws ban valuable inquiry—sometimes putatively morally valuable, but sometimes purely zetetically valuable. Others grant that experiments prohibited under Directive 2010/63/EU are zetetically right but all-things-considered wrong.

Second, on this view, what makes the zetetic evaluation a distinct subject matter is a restriction, not on the flavor of norms or value, but on the type of acts. Zetetic evaluation only evaluates acts of inquiry. But it’s hard to see how this is a distinctive class of act. Gathering evidence is often offered up as a canonical act of inquiry. But everything we do involves our gathering evidence: for example, when I ate tacos today, I acquired evidence that my salsa had not gone bad. If eating tacos is thereby a form of inquiry, then anything is.\(^{11}\) If so, then either there are distinctively zetetic norms or the literature on the zetetic has no subject matter.

---

\(^9\) For a recent dissenting view, see Rinard (2017).

\(^{10}\) This view has been defended extensively by Thorstad (manuscript a,m).

\(^{11}\) I discuss this question further in $5.1.$
1.2 Consequentialisms

*Epistemic consequentialism* holds that the epistemic rational total belief state\(^{12}\) for an agent is the belief state that is most conducive to promoting epistemic value. Epistemic value is often identified with accuracy; the most widely accepted form of epistemic consequentialism is alethic consequentialism. Epistemic rationality does not, of course, require omniscience; we are not required to maximize actual accuracy. Instead, consequentialists hold, we ought to do the best we can in promoting accuracy in our beliefs, given our limited information. Some hold that rationality requires avoiding having accuracy-dominated beliefs: beliefs that are guaranteed to be farther from the truth than some alternative possible beliefs, no matter what the world is like. Some hold that rationality requires holding beliefs that maximize expected accuracy. Still others hold that rationality requires that beliefs be reliably formed.

Zetetic consequentialism is a related view, describing norms of inquiry. And a natural version of this view is alethic consequentialism, according to which the value that zetetic norms require promoting is accuracy. And so zetetic and epistemic versions of alethic consequentialism fit neatly together. Epistemic consequentialists adapt pragmatic consequentialism, and its decision theories, to the epistemic case. Decision theories take as input decision problems, characterized in terms of states of the world, available acts, values, and states of uncertainty. Epistemic alethic consequentialism restricts the form of value to specifically epistemic value (identified with accuracy), and the possible set of acts under evaluation to “epistemic acts”.\(^{13}\) Zetetic alethic consequentialists also restrict the form of value to epistemic value. But all actions, not just epistemic acts, can be zetetically evaluated. This allows for the zetetic evaluation of actions like conducting specific experiments, collecting evidence, or earning PhDs. Zetetic alethic consequentialism, then, holds that the zetetically best actions are those that are most conducive to promoting *epistemic* value.\(^{14}\)

---

\(^{12}\) Beliefs/suspicions/rejections, credence functions, credal representor, comparative confidence preorder, etc.

\(^{13}\) Interpreted in a non-voluntarist way.

\(^{14}\) It doesn’t follow, however, that epistemic consequentialism is a special case of zetetic consequentialism. It’s generally thought that epistemic norms are agent-relative: *my* epistemic rationality is not influenced by *your* rationality or accuracy, or whether I help to improve them. But the norms of inquiry may be agent-neutral. This hypothesis is most plausible in the case of scientific inquiry, where, for example, division of cognitive labor may support some agents pursuing inquiry that is less likely to improve their own accuracy but more likely to contribute to the accuracy of the greater scientific community. See Kuhn (1977); Kitcher (1990) for related discussion.
2 Case Study

2.1 Observation vs. intervention

We’ll first consider a problem case for zetetic consequentialism and then adapt it for epistemic consequentialism.

Microecosystem Case (zetetic variant).

You’re a biologist and have encountered an isolated microecosystem with its own peculiar flora and fauna, which will change over time in unpredictable ways: various populations might overtake others, suffice resource scarcity, face new infectious diseases, etc. So if you observe the microecosystem over time with minimal intervention (call this option “observe”), you’ll face a great deal of uncertainty about its future, various of its hidden features, etc. But you have another option: you can set the microecosystem on fire, reducing it to ash, in such a way that it’ll predictably remain nothing but ash forever after. (Call this option “intervene”.) Doing so, your present and future beliefs about the microecosystem will almost certainly be far more accurate than they would be if you merely observed.

Thesis: you are not required by the norms of inquiry to set the microecosystem on fire. More generally, setting the microecosystem on fire is not a zetetically better action than merely observing it without intervening.

So, the norms of inquiry permit actions that predictably generate less accurate beliefs than other available actions. That means that alethic consequentialism about inquiry is false.

Next we’ll adapt the example for the epistemic case:

Microecosystem Case (epistemic variant).

Same as before, but now we are considering possible epistemic acts. You know that if you believe that you’ll set the microecosystem on fire (call this option “intervene$_e$”), then that belief will cause you to form the intention to do so, which in turn will cause you to do so. If you don’t believe it (call this option “observe$_e$”), then you won’t be caused to do so.
Intervening will predictably cause your present and future beliefs about the microecosystem to be more accurate.

Thesis: you are not required by the norms of epistemic rationality to believe you’ll set the microecosystem on fire. More generally, believing you’ll set the microecosystem on fire is not an epistemically better action than believing you’ll observe it without intervening.

So, the norms of epistemic rationality permit epistemic acts that predictably generate less accurate beliefs than other available epistemic acts. That means that alethic consequentialism about epistemically rational belief is false.

2.2 Objections and modifications

Objection #1. Entropy. Why should we accept that, in the Microecosystem case, your beliefs will be more accurate if you choose intervene? After all, the ashen state is more disorganized, so more information will be necessary to accurately describe it. (That is, its informational entropy, like its thermodynamic entropy, is higher than that of the microecosystem.) So there are plenty of propositions about the ashes of the demolished microecosystem that you will not be able to accurately predict.

Reply. If you interpret the case this way, tweak the example: perhaps you have the option of converting the entire ecosystem into pure diamond, highly predictable in its well-organized atomic structure and stability under extreme pressures and temperatures. House the diamond in lead-lined concrete; store it deep underground, where it’s highly likely that no one will disturb it. Reassign “intervene” to refer to this option.

Objection #2. Counterfactuals. Intervening may make your beliefs more accurate about nonmodal facts about the microecosystem, but your future beliefs will be less accurate about many modal propositions: propositions about the different ways the microecosystem could have evolved over time had you not set it on fire. If you observe, you’ll learn answers to many of these questions; if you intervene, you won’t. And there’s an asymmetry here: if you observe, your beliefs will still be accurate about counterfactuals involving how the system would have evolved if you’d intervened, since by hypothesis, these are predictable ex ante.
Reply. If you interpret the case this way, tweak the example: you know in advance that if you merely observe the microecosystem, there’ll be many facts about it that you’ll never learn: it will be complex enough that many parts will be unobservable. It’ll persist long after your death, and its future thereafter will be similarly random and unpredictable. The scenario can be refined to one in which the modal accuracy you hope to gain in choosing observe is more than outweighed by the nonmodal accuracy you can confidently expect to gain from choosing intervene.

Objection #3. Greater ramifications. observing doesn’t only affect the accuracy of our beliefs about the system. It has the potential to provide illuminating evidence that bears on our general understanding of biology (physics, etc.) If our beliefs become more accurate about general facts like scientific laws, this will improve the accuracy of vast swathes of our beliefs about particular propositions entailed by those laws together with our evidence. And so observing may still maximize expected accuracy.

Moreover, these propositions may be more epistemically important. And so their contribution to the global accuracy of our total belief states might be weighted more heavily than unimportant, local facts, like how many blades of grass are in the microecosystem.

Reply. If you interpret the case this way, tweak the example: there are lots and lots of microecosystems like this, all internally unpredictable in how they’ll evolve over time. If the system you’re considering intervening on has anything important to reveal about more general scientific laws, it’s highly likely that one of the other microecosystems will reveal that same information independently. But if you intervene, then at least you can be sure of this microecosystem’s future and its hidden depths.

With these refinements, of course, one might think that it’s not so urgent that you not intervene. But remember, the alethic consequentialist doesn’t just say that you’re permitted to intervene; it says you’re

---

15 Or, if the relevant accuracy is not agent-relative, everyone’s death.

16 A frustrating challenge for both sides: if we aim to include all potentially relevant propositions in the weighting, there are infinitely many propositions that are in principle learnable only if you observe, and infinitely many that are learnable only if you intervene. And there’s no reason to think either of these infinite sets is greater than the other. For epistemic consequentialism, then, it’s often assumed that the accuracy of one’s total belief state is relativized to a finite agenda of propositions. While we can generalize to the infinite case, doing so makes it difficult to formulate thought experiments that we can form intuitive, pretheoretic judgments about.

But suppose we stipulate some finite set of propositions that are on the inquiry agenda for our microecosystem. For some selections of agendas, observing will have greater expected accuracy; for others, intervening will. To vindicate the idea that observing maximizes expected accuracy, one would need to justify prohibiting the second sort of agenda. It’s hard to see how the alethic consequentialist could justify this prohibition, though. Propositions outside of the agenda can be true or false. No purely accuracy-based consideration will justify treating their truth as irrelevant.
required to! And this is implausible: the norms of belief and inquiry do not require you to intervene to make the microecosystem predictable.

3 Diagnosis

3.1 Discovering vs. creating truth

Thesis: The aim of belief and inquiry isn’t to have true (or accurate) beliefs as such. There are two ways to end up with true beliefs: by discovering the truth, or by creating the truth. Inquiry, and epistemic rationality, don’t aim to create truths. Indeed, as I’ll argue in §4, inquiry positively aims not to create truths, at least in many cases. The aim of inquiry is to discover truth—truths already, in some sense, out there.

I think this thesis is simple and intuitively obvious. And it provides a straightforward diagnosis for why intervene is not zetetically or epistemically required, and may even be zetetically and epistemically prohibited. Creating truths can make it impossible to discover truths.

There are some subtleties to consider. First, "discover" sounds binary. It may be better to claim that aim of inquiry is to increase our gradational accuracy by bringing our beliefs or credences more in line with the world as we find it.

Second, the distinction between discovering the truth and creating the truth is not clearcut. There are acts that raise the chances of some truths without guaranteeing them, and indeed, there are acts that may lead to discovering truths but may also risk creating truths:

Urn.

I’ll flip a coin. If the coin lands heads, I’ll pull a ball out of an urn full of white and blue balls and discover its color. If the coin lands tails, I’ll pour a bunch of blue dye into the urn, and make it true that the next ball I pull from the urn will be blue.

Dumplings.

I’ll ask my partner whether they want dumplings. There’s some chance that from their answer, I’ll discover their preferences. But they might instead infer that I want dumplings,
and insofar as they prefer for my preferences to be satisfied, I might inadvertently (but predictably) change their preferences so that they come to want dumplings. In both cases, there’s a risk of creating the truths that I hope to discover.

But even if the distinction between discovering and creating truths is blurry, that doesn’t mean the aim of inquiry is blurry. It might be hard to zetetically evaluate a borderline case of discovering/creating truth in how well it lives up to inquiry’s aims, but that doesn’t mean it’s indeterminate whether inquiry also aims at creating truths.\textsuperscript{17}

Third, “already out there” shouldn’t be taken to mean that the relevant truths must concern the present or past. And indeed, “out there” shouldn’t be taken to mean that the relevant truths concern the inhuman world, or even the world beyond the inquirer. We may ourselves be subjects of inquiry. But when we are, we don’t satisfy zetetic norms by simply making ourselves conform to our beliefs about ourselves, thereby ensuring that those beliefs are true. Of course, we must make decisions, and sometimes the decisions we make answer questions we previously had. But zetetic norms do not govern these decisions in ways that require conforming the decisions to our predictions, even when doing so generates the greatest accuracy.

3.2 \textit{Relations between epistemic and zetetic consequentialism}

The problem for alethic consequentialism stems from the fact that, intuitively, epistemic and zetetic norms have a mind-to-world direction of fit. The aim is to take epistemic and pragmatic acts that allow us to conform our minds to the world. But alethic consequentialism holds that epistemically and zetetically good acts are those that conduce to our having beliefs that match the truth as closely as possible. Having beliefs that match the truth is \textit{neutral} about direction of fit: it can equally well be achieved by shaping the world to conform to our beliefs.\textsuperscript{18}

This diagnosis cuts off a \textit{prima facie} plausible conception of the relation between epistemic and

\textsuperscript{17} Compare: at my local bar, I request their hoppiest, sweetest beer. I value hoppiness and disvalue dryness. The bartender has nothing ideal to offer: only a dry, hoppy IPA and a sweet, non-hoppy Belgian ale. It may be indeterminate which of these better satisfies my requirements, but that doesn’t mean that it’s indeterminate whether I value non-hoppiness or disvalue sweetness: I determinately do neither.

\textsuperscript{18} See Berker (2013b); Carr (2017); Konek & Levinstein (2019) for discussion.
zetetic norms: that zetetic, but not epistemic, norms are alethic consequentialist. It’s worth considering why this hypothesis is tempting, and why cases like Microecosystem show that it won’t work.

Epistemic nonconsequentialism is often motivated by examples involving epistemic tradeoffs, where one can take on a belief that one is in a position to know is false, in order to predictably gain enough true beliefs to counterbalance the falsehood.\textsuperscript{19} It is also motivated by cases of self-verifying beliefs, where epistemic consequentialism requires agents to adopt beliefs that will cause themselves to be true.\textsuperscript{20} Such cases relate closely with self-falsifying beliefs, where the beliefs that are most accuracy-conducive are often incoherent.\textsuperscript{21} But since the latter generate a clash between plausible epistemic norms—accuracy vs. coherence—they are sometimes taken to generate epistemic dilemmas. Self-verifying beliefs needn’t generate any similar clash. Instead, they shine a spotlight on the fundamental problem of alethic consequentialism: that it ignores direction of fit. Beliefs have various causal effects in the world, and sometimes probabilify themselves. So epistemic consequentialism sometimes requires us to mold the world to our beliefs. The Microecosystem cases are an example.

In the case of epistemic tradeoffs, alethic consequentialists often bite the bullet, sometimes offering accompanying error theories of our intuitions.\textsuperscript{22} In the case of self-verifying beliefs, some reject that there’s a bullet to bite. Many insisting that it’s intuitively irrational to turn down a “free true belief”, and so hold that these examples support alethic consequentialism.

There is a response available to the epistemic nonconsequentialist. Yes, there is some appeal to the idea that you oughtn’t turn down free true beliefs. But the relevant norm ought to be thought of as pragmatic, not epistemic. Why? Suppose you happen to desire having the most accurate beliefs possible. So as a matter of rational desire-satisfaction, you \textit{pragmatically} ought to accept epistemic tradeoffs, self-verifying beliefs, and incoherent beliefs, whenever doing so maximizes expected accuracy. In doing so, though, you might end up with beliefs that are \textit{epistemically} irrational. Epistemic norms are objective;

\begin{enumerate}
\item Berker (2013a); for other examples of epistemic tradeoffs, see Firth (1978/1998); Greaves (2013). Note that the epistemic Microecosystem case does not involve an epistemic trade-off, since the belief that you’ll set fire to the microecosystem is self-verifying.
\item See Berker (2013b); Carr (2017).
\item If I know that $p$ is true if and only if I don’t believe $p$, then the most accurate belief state will involve doxastic neutrality on $p$ and full-throated rejection of $\neg p$. But this state is incoherent: if I fully reject $\neg p$, then I ought to fully believe $p$. See Caie (2013); Greaves (2013).
\item E.g. Pettigrew (2018); Singer (2018).
\end{enumerate}
your subjective desires are epistemically irrelevant.

Notice that this affects how we conceive of the epistemic/pragmatic distinction. If the above is correct, we cannot think of the distinction as determined either by the domain of objects of evaluation (belief states vs. actions) or by the relevant axiology (accuracy vs. moral value or utility). First, the nonconsequentialist explanation above assumes that there can be pragmatic norms on belief, not just action. This is independently plausible: bigoted beliefs can be morally wrong, not just epistemically; motivated reasoning can be preferable to pure epistemic rationality. Second, the nonconsequentialist explanation holds that accuracy can be pragmatically, not just epistemically, valuable. For agents who care about accuracy, its pursuit is pragmatically rational. For agents who don’t, the value of accuracy only affects epistemic evaluation.

Where the pursuit of accuracy is pragmatically rational, we can assess not just beliefs but also actions for accuracy-conduciveness. And so, you might think, we’ve moved from epistemic to zetetic norms. Even nonconsequentialists about the epistemic can still be consequentialists about pragmatic instrumental rationality. And prima facie, it’s initially appealing to characterize zetetic norms within pragmatic instrumental rationality: as norms for the effective pursuit of accuracy. In other words, it’s initially appealing to assume that zetetic norms are alethic consequentialist norms.

If so, then there can be conflicts between epistemic and zetetic norms: the latter, but not the former, are alethic consequentialist. And prima facie, this seems to offer a tidy explanation of the split intuitions on epistemic tradeoffs and self-verifying beliefs. The epistemic nonconsequentialist can agree with the consequentialist that there’s an important sense in which epistemic tradeoffs and self-verifying beliefs are obligatory. Moreover, the norms requiring these attitudes are intimately connected with epistemic norms: the norms of inquiry are surely closely related with the epistemic. But occasionally they’ll conflict. Zetetic norms, on the consequentialist picture, require choosing epistemic tradeoffs that are epistemically impermissible, and require adopting self-verifying beliefs that are epistemically unrequired.

This explanation has a lot going for it: for example, it seems to make sense of a form of epistemic trade-off that many find compelling: cases where scientists incorporate false idealizations into theories or models. These idealizations promote various goals in scientific inquiry: improving tractability, characterizing specific phenomena in isolation from complicating extraneous factors, etc. It’s tempting to
cite these idealizations as virtuous epistemic tradeoffs: cases where one adopts some false beliefs as a means of acquiring a larger number of true beliefs.

But there's a problem. The zetetic Microecosystem case shows that norms of inquiry are not alethic consequentialist norms. In scientific inquiry, we are not required to change the world in such a way as to make it fit our beliefs. Inquiry might require altering the world to make it more knowable to us—in ways that expose its truths. But it doesn’t require altering the world to make it more predictable to us—in ways that make its truths conform to our predictions.

So we should reject alethic consequentialism about both zetetic and epistemic norms. Even if these norms diverge in various respects, and even sometimes conflict, both aim at discovering, rather than creating, truths.

4 Is intervention a zetetically evil?

The previous section argued that zetetic norms don’t require don’t require intervening into the world make it conform to our beliefs, even when doing so uniquely maximizes the expected accuracy of our present or future beliefs. But a stronger claim is plausible: that such interventions violate zetetic ideals, or are (sometimes or always) zetetically impermissible. Before, we saw that such interventions don’t contribute positive zetetic value; now, we’ll explore the idea that they actively decrease zetetic value.

More precisely, we’ll explore the hypothesis that interventions into a question under inquiry are zetetically disvaluable. This involves two specialized notions: a question under inquiry and an intervention into a question.

When an agent engages in inquiry, she aims to discover the answer to some question: call this the question under inquiry. I’ll represent a question as a partition, \( Q = \{q_1, \ldots, q_n\} \). Each \( q_i \) is a possible answer to \( Q \).

An intervention into \( Q \) is an act that causally alters the objective probabilities of some answers to \( Q \). In the simplest case, one can intervene into \( Q \) by making some \( q_i \) in \( Q \) true. But one can also intervene in ways that only raise (or lower) the chances of some \( q_i \). Note that whether some action counts as an

\[^{23}\text{Why objective probabilities? In mere observations of the world, we causally affect our own subjective probabilities of answers to our questions. But this still involves our belief states coming into conformity with the world, rather than the reverse.}\]
“intervention” is question-relative: an act may constitute an intervention into \( Q \) and not an intervention into \( Q' \). Indeed, an intervention into \( Q \) might be zetetically ideal for inquiry into \( Q' \)?

Why think that intervention into a question under inquiry might be zetetically bad? Consider an example, quite similar to Microecosystem:

\[birds\]

You’re inquiring into the migratory patterns of a group of birds. The question under inquiry is where these birds will be next January; call this \( Q_{birds} \). Their migratory patterns seem complex and unpredictable. You don’t have the means to track the birds. What you can do is lure these birds into your yard, trap them, and keep them in your basement. Then, and only then, you can predict with great accuracy the answer to \( Q_{birds} \): next January, they’ll be in your basement.

It’s clear that the norms of inquiry don’t require trapping the birds. But I want to push a further intuition: that in this example, trapping the birds is zetetically worse than, say, merely observing the birds.\(^{25} \) If zetetic norms are maximizing norms—if they require choosing one’s very best options—then trapping the birds is impermissible. Intervening to force a migratory pattern on the birds—one whereby the birds can only migrate within my basement—undermines the aims of inquiry.

Notice: the claim under consideration is not that any form of intervention in the context of inquiry is all-things-considered impermissible. We can have plenty of nonzetetic reasons to intervene in questions under inquiry. Zetetic evaluation is one form of evaluation among many; it can be outweighed.

And the claim is not that any form of intervention is zetetically impermissible. In various ways, intervention is often inevitable. Whenever we take actions beyond bare observation—opening the fridge, performing experiments—we intervene in the world in significant ways. Typically, however, these are not interventions into questions under inquiry.

But in fact, the claim is not even that any form of intervention into questions under inquiry is impermissible. Our actions can have significant and unavoidable effects on the very propositions we’re

\(^{24}\) Note: I don’t assume that at most one question can be actively under inquiry at a time.

\(^{25}\) Trapping the birds is also, of course, morally impermissible, for separate reasons.
investigating. On the popular Copenhagen interpretation of quantum mechanics, for example, acts of observation often unavoidably cause indeterminate properties (position, momentum, etc.) of particles in superpositions to “collapse” into determinate properties. More banally, when I inquire into whether my silent bedmate is sleeping, all of my options for investigating the answer raise the probability that I will find them awake.

Instead, the claim under consideration is that interventions into questions under inquiry are zetetically disvaluable; if an action intervenes on a question under inquiry, that’s a pro tanto zetetic reason against the action. The Microecosystem and Birds cases suggest that this is so. It’s unclear, however, whether this is always so, or whether there are exceptions.

Objection #4. Sometimes we inquire into our own future actions. (Will I give into temptation and have a beer tonight, even though I’m trying to cut down? Will I get into another argument with my partner over pandemic precautions?) We have no choice but to affect which answer turns out to be correct. So interventions here cannot be zetetically bad.

Reply. A variety of responses are available. First, one might allow that because intervention is unavoidable, this is a case where it’s zetetically permissible, but still insist that such interventions are zetetically suboptimal. (It’s just that the optimal outcome is unattainable.) Alternatively, one might hold that questions that are under the agent’s immediate control at a time \( t \) are not possible or permissible questions under inquiry at \( t \). (I’ll discuss this hypothesis further in \$5.1.) Finally, one might think: inquiry into how one will act in the future ends before one acts. Between the two, there’s a period of deliberation.\(^{26}\) Distinguish between inquiring into how one will act and deliberating about how to act. Very often, inquiring into how you’ll act shades into deliberating on how to act, particularly once you begin considering one’s possible reasons for your future action.\(^{27}\) Once you’re deliberating on how to act, you’ve shifted the question under inquiry: the question is no longer how you will act, but how you should act. In that case, norms, rather than your own future actions, are the subject of inquiry. But then, by ultimately acting, you don’t intervene on the new question under inquiry: you don’t affect what you should do.

---

\(^{26}\) I use the generic formulation to refrain from commitments on whether this is universal or merely typical.

\(^{27}\) This might explain why natural examples of inquiry into how you will behave in the near future involve irrational actions. These examples isolate the non-normative question of how you will act from the confound of related normative questions.
Objection #5. Suppose I’m inquiring into the question: what would happen if I were to perform some action φ? In order to learn the answer to this question, I can’t sit back and inertly observe; I must take action and φ! So surely some inquiries permit interventions into the world—indeed, require interventions.

Reply. Where Q is the question what would happen if I were to φ?, its answers are propositions expressed by subjunctive conditionals of the form: were I to φ, then outcome o₁ would obtain; were I to φ, then o₂ would obtain; etc. When you φ, you raise the objective probability that you φ (to 1). But that’s not one of the answers to the question under inquiry. In order to intervene on the question under inquiry, you would need to affect the objective probability of some of the subjunctive conditionals (the answers to Q); but your φing doesn’t do that.

In §2 and §3, I argued against the idea that zetetic norms could be explained teleologically as aiming at maximizing the accuracy of our beliefs. I haven’t provided a positive theory of zetetic norms. So that leaves an open question: what, if anything, is zetetically wrong with intervening into questions under inquiry?

5 What’s wrong with intervention?

If I’m right that intervention into questions under inquiry is (at least sometimes) zetetically disvaluable, this fact stands in need of explanation. If interventions were merely not valuable, then they might simply be irrelevant to inquiry. They’d have no effect on zetetic value, any more than the number of raspberries in my refrigerator has any effect on the value of my present philosophical inquiry. But if interventions into questions under inquiry can be actively bad for inquiry, then a complete understanding of the norms of inquiry will require understanding why that’s so.

I pose this as a puzzle. This paper won’t settle on how to resolve it. Instead, I’ll explore some potentially promising answers, while also noting some reasons for hesitancy about each.

5.1 Explanation #1: Inquiry and intervention are incompatible actss

Explanation #1: zetetic norms are question-relative, and place high value on active inquiry into the question under inquiry. Intervening into Q is incompatible with actively inquiring into Q. Interventions
into $Q$ involve actions, and actions are the result of decisions. Inquiry into $Q$ ends before decisions, at the stage of deliberation. Inquiry into $Q$ and deliberation about whether to affect its answers are distinct and incompatible activities, governed by distinct norms. When you are inquiring into $Q$, you occupy a third-personal perspective, guided by evidence. When you are deliberating about whether to affect the answer to $Q$, you occupy a first-personal perspective, guided by pragmatic considerations. Once you start deliberating on whether to intervene into $Q$, you've cut off inquiry into $Q$. In this sense, deliberation crowds out inquiry.\footnote{There's a relation between the problems for zetetic consequentialism I've discussed and the literature on the "deliberation crowds out prediction" (DCOP) thesis, defended in Spohn (1977); Levi (2007); Price (2007). Motivating examples for DCOP often involve self-frustrating decisions (e.g., the familiar Death in Damascus case from Gibbard & Harper (1978)). Self-gratifying cases are generally less concerning for practical decision theory, but in the epistemic case, self-verifying cases are arguably just as puzzling (see especially Carr (2017)). Note that unlike proponents of DCOP, this paper is not committed to the claim that in deliberative contexts, it's irrational to hold any credence in propositions about which act one will choose.}

Explanation #1 faces a few challenges. First, it presupposes that there is a distinctive set of activities that constitute, or are compatible with, "inquiring into $Q". What activities fall into this category? The answer is almost certainly both complicated and vague. When I'm inquiring into $Q$, $Q$ itself need not constantly occupy my conscious thoughts. (In inquiring into when I'd receive my COVID-19 vaccination, my mind was often occupied by thoughts about whether the relevant information will come from my health-care provider's website, my insurance company, or my governor, whether university professors are counted among "teachers", etc.: my stream of conscious thought was not constantly repeating: "When will I get vaccinated?")

Instead, one might think that various different kinds of activities aimed at discovering the answer to $Q$ can involve inquiring into $Q$, and shifting attention to the means to this end is compatible with engaging in inquiry. Indeed, it might be that inquiry is sometimes dispositional rather than occurrent. Just as one can dispositionally believe $p$ while consciously thinking about other, unrelated things, perhaps one might dispositionally inquire into $Q$ even when one is engaged in entirely unrelated activities.

But now, we should ask: why isn't intervening into $Q$ one of the activities that are compatible with inquiring into $Q$? Indeed, why isn't it a possible means to the end of discovering the answer to $Q$? There may well be a satisfactory answer to this question that's consistent with Explanation #1, but I don't know what it is.
Second, recall that intervention can be stochastic: rather than definitely making some answer to Q true, one might merely raise its objective probability. Recall the Urn example in §3.1. The relevant act—deciding whether to draw a ball from an urn or dye all the balls in the urn on the basis of a coin flip—constitutes an intervention into the question under inquiry. I raise the higher-order objective chance (50%) that I’ll raise the objective chance of pulling out a blue ball to 1. The present objective chances of the future objective chance of p determines the present objective chance of p. So this act constitutes an intervention. But there’s a 50% chance that it will result in a canonical act of inquiry: observation.

Third and relatedly: arguably, unknowing interventions into a question under inquiry are zetetically disvaluable (even if zetetically blameless). But it’s hard to see how these are incompatible with inquiry: why couldn’t some acts of inquiry have unforeseen downstream consequences that constitute interventions into the question under inquiry? In many cases, there’s some, perhaps quite low, probability that canonical acts of inquiry—gathering evidence, conducting experiments—will result in intervention into the question under inquiry. But that means that these canonical acts of inquiry count as interventions: they affect objective chances of objective chances of answers to questions under inquiry.

Finally, this suggests that the problem with intervening is merely that it interferes with a zetetically valuable activity: inquiring. In that case, the problem with intervening in the Microecosystem case is that it’s a distraction—it’s zetetically problematic in the same way, and to the same extent, that playing videogames is. But intuitively, setting the Microecosystem on fire is zetetically worse—perhaps even much worse—than playing videogames. And the hypothesis currently under consideration offers no explanation for that.

None of these worries makes for a decisive objection. After all, we allowed that interventions into questions under inquiry may not always be zetetically disvaluable. So it may be that such interventions are only nonideal insofar as they interfere with the activity of inquiring. And indeed, it might be that intervening into the Microecosystem is zetetically worse than playing videogames because the former, and not the latter, interferes in future inquiries. But more needs to be said about why this would be so.
5.2 Explanation #2: Only some propositions are relevant to zetetic value.

*Explanation #2:* the zetetic aim is accuracy only with respect to a specific set of propositions: propositions that are causally independent of the agent’s present available actions.

Call the set of propositions an agent is capable of entertaining \( \mathcal{P} \). On the present hypothesis, there’s some set \( \mathcal{A}_t \subseteq \mathcal{P} \) of zetetically significant propositions at a time \( t \). If at \( t \), the agent can affect the objective chance of \( p \), then \( p \in \mathcal{P} \setminus \mathcal{A}_t \), and at \( t \), future gains in accuracy about \( p \) are zetetically valueless. Only improvements in accuracy with respect to members of \( \mathcal{A}_t \) are zetetically valuable.

A variety of propositions are members of \( \mathcal{A}_t \): among them, propositions about events that take place before \( t \), as well as propositions that are noncontingent, either necessarily true or necessarily false. Various future contingent propositions are also in this category: there are many possible future events that an agent cannot affect the objective chances of (e.g., in Minkowski spacetime, events outside of the agent’s future light cone).

But what about cases like Microecosystem and Birds, where an agent is inquiring into questions whose answers are not causally independent of the agent’s present available actions? Is it zetetically valueless to improve one’s accuracy with respect to these questions? According to Explanation #2: yes! But that doesn’t mean that the inquiry into these questions is valueless. Instead, it’s derived from improvements in accuracy about other propositions.

Suppose an agent is inquiring into \( Q \) and the members of \( Q \) are not members of \( \mathcal{A}_t \). The agent can improve her accuracy with respect to a related set of propositions in \( \mathcal{A}_t \): propositions about how answers to \( Q \) do or do not depend on the agent’s available acts. We can represent these as non-backtracking counterfactual conditionals: where \( \Phi \) is a partition of propositions characterizing available acts for the agent at \( t \), we’ll consider counterfactuals of the form \( \Gamma \phi_j \implies q_i \), where \( \phi_j \in \Phi \) and \( q_i \in Q \).\(^{29}\) Call these “dependency hypotheses.”\(^{30}\) These propositions are members of \( \mathcal{A}_t \), and so on this hypothesis, are zetetically significant. Acts that improve the accuracy of the agent’s present or future beliefs in these propositions are zetetically better than those that don’t.\(^{31}\)

\(^{29}\) For readability, I’ll use \( \phi_j \) and \( q_i \) as variables both for propositions and sentences expressing those propositions.

\(^{30}\) The machinery of dependency hypotheses is developed in Lewis’s (1981) causal decision theory. The dependency hypotheses I use are partial; a complete dependency hypothesis is a maximal consistent conjunction of partial dependency hypotheses.

\(^{31}\) Notice: dependency hypotheses aren’t the only way of representing the relevant form of causal structure. We could instead
One might think this makes interventions into questions under inquiry zetetically irrelevant, rather than zetetically disvaluable. But they aren’t: intervening into questions under inquiry will often make it harder for the agent to improve her future accuracy with respect to some dependency hypotheses.

Return to the Microecosystem case. Suppose you’re inquiring into some $Q$ about the future of the microecosystem: for example, how many plant species will it contain at some future time. If you intervene, you might make it harder to improve your future accuracy about propositions of the form $\diamondsuit \text{observe } q_i \downarrow$ for various $q_i \in Q$. And you gain little or no accuracy about propositions of the form $\diamondsuit \text{intervene } q_i \downarrow$ to compensate. After all, ex hypothesi, before you choose an act, you’re in a position to accurately predict that if you were to intervene, the number of plant species in the microecosystem would be 0. You can expect to gain little or no accuracy on that proposition by intervening rather than observing: your degree of accuracy in propositions about what would happen if you intervene will be high either way. But what about propositions of the form $\diamondsuit \text{observe } q_i \downarrow$? While you have a significant chance of improving your accuracy about these propositions in the future if you observe, intervening blocks you from improving your accuracy about these.

Another example: consider a simplified version of the Urn case above. Your options are draw and dye, and the question under inquiry is: what color will the next ball drawn be? You can predict that whatever you choose, you’ll end up with accurate beliefs about $\diamondsuit \text{dye } q_i \downarrow \text{the ball will be } x \downarrow$, for any color $x$: accurate high credence if $x$ is blue, and accurate low credence if $x$ is any other color. But whereas you’ll learn the answers to $\diamondsuit \text{draw } q_i \downarrow \text{the ball will be } x \downarrow$ if you draw, you won’t if you dye.

Explanation #2 helps to explain why some interventions into a system are zetetically valuable. In the Microecosystem case, for example, it might be zetetically valuable to introduce a new species into the Microecosystem to see how it interacts with other species and with abiotic components of the ecosystem. This form of intervention helps us gain accuracy about a variety of questions: questions about what would happen if we introduced that species into the environment, and perhaps also questions about various properties of the Microecosystem before the introduction of the new species. Notice, though, that these are questions whose answers we can’t presently control: the former question’s answers are dependency hypotheses, and the latter’s answers are propositions about the past. In other words, this is an

---

represent this information using, e.g., causal models such as causal Bayes nets (Pearl, 2000; Spirtes et al., 2000).
intervention into the *microecosystem*, but not necessarily an intervention into the *question under inquiry*. And the zetetic benefits it affords are related to learning truths that are causally independent of our actions. Of course, we also learn the answer to the question: *will this particular new species ever find its way into the microecosystem?*. But because we learn this merely by making one of its answers true, that alone contributes no zetetic value.\(^{32}\)

If Explanation \#2 is correct, then we can expect that intervention into a question under inquiry is only sometimes zetetically disvaluable. There may well be cases where such interventions can be predicted to improve your future accuracy about the relevant dependency hypotheses.

Notice that this hypothesis involves a view almost identical to alethic consequentialism: zetetic value derives from accurate beliefs, but only with respect to a restricted set of propositions, \(\mathcal{A}_t\). So one might think: if this is the correct explanation, then alethic consequentialism was correct in spirit, if not in letter. The correct view is still consequentialist and the source of value is still accuracy.

The reality is more complex, though. First, even if this is the correct theory of zetetic norms, it won’t work for epistemic norms. After all, if your accuracy with respect to propositions outside \(\mathcal{A}_t\) are not sources of *epistemic* value, then the resulting form of consequentialism won’t be able to explain why incoherence in our attitudes toward these propositions is epistemically irrational.

One might think that we can recover coherence norms for such propositions by looking at disjunctions or conjunctions of dependency hypotheses. After all, it’s often accepted that \(\Gamma p \implies q \land \neg p \implies q'\) is equivalent to \(q\).\(^{33}\) So even if \(q\) isn’t in \(\mathcal{A}_t\), we may be able to recover accuracy-based constraints on \(q\) from members of \(\mathcal{A}_t\). But even if the conjuncts of \(\Gamma p \implies q \land \neg p \implies q'\) are members of \(\mathcal{A}_t\), their conjunction might not be. After all, if \(q\) is under the agent’s control, so is any proposition equivalent to it. And so accuracy-based considerations governing beliefs about the conjuncts won’t require them to cohere in epistemically rational ways with beliefs about the conjunction.

Epistemic norms govern *all* beliefs—not just beliefs about propositions we have no control over. Given this, we can’t use Explanation \#2 to explain why intervening \(e\) is *epistemically* unrequired in the Microecosystem case. I suspect that the most likely explanation will be that epistemic norms are gen-

\(^{32}\) Thanks to [omitted] for the example.

\(^{33}\) Assuming a strong centering constraint on counterfactuals.
ually nonconsequentialist.

So Explanation #2 has a potential cost: it means that epistemic and zetetic norms are surprisingly disunified, in their axiology or deontic theory or both.

Does Explanation #2 at least support the idea that a restricted form of alethic consequentialism is true in the zetetic case? It’s not obvious whether this should really count as a form of alethic consequentialism.

First, it’s not purely alethic. We need some explanation for why accuracy about only some propositions counts. This explanation can’t simply appeal to the intuitive value of truth as a property of beliefs: after all, some true beliefs are zetetically valueless.\(^{34}\)

Second, it might not even be consequentialist. Determining what precisely counts as a consequentialist theory is a can of worms.\(^{35}\) It’s often thought that a necessary condition on consequentialism is that it treats value as agent-neutral (e.g., McNaughton & Rawling 1991; Howard-Snyder 1994; Pettit 1997; Brown 2011).\(^ {36}\) Given Explanation #2, zetetic value is agent- and time-relative: these parameters determine which propositions are zetetically relevant. So it may well be disqualified from consequentialism.\(^ {37}\)

6 Conclusion

One might think that the ideal of “merely observing” holds only in the case of scientific inquiry. There, it’s plausible enough that we aim to learn about the world out there that’s independent of our meddling. After all, one might reason, in scientific inquiry, the questions under inquiry are questions about the very general causal structure of the world—questions whose answers our interventions, if even possible, are likely to obfuscate rather than illuminate. In mundane inquiry, one might think, there’s no similar constraint on the objects of inquiry.

But in fact, mundane inquiry is also undermined by intervention into the question under inquiry. If

---

\(^{34}\) Furthermore, the explanation for why some propositions are zetetically insignificant may not be consequentialist.

\(^{35}\) The literature on consequentializing moral theories attests to this. See, e.g., Portmore (2007); Brown (2011).

\(^{36}\) For dissent see, e.g., Dreier (1993).

\(^{37}\) It’s worth noting that, by this standard, some views often called “epistemic consequentialism” may not count as consequentialist either; see footnote 14.
I’m inquiring into my partner’s preferences re dumplings, it’s zetetically better to word the question in such a way as to avoid influencing their preferences to match mine, even if influencing their preferences would result in more accurate beliefs (and a more enjoyable dinner for me). If I’m inquiring into whether the dumplings will taste good, it’s zetetically better to taste them than to douse them with sweet chili sauce. Even if I know sweet chili sauce will make them taste good, dousing them doesn’t help me to discover the truth. It just replaces the truth.

This observation reveals important features of zetetic and epistemic normativity, but also leaves open serious questions. It reveals that alethic consequentialism, despite its seductive simplicity and apparent explanatory power, cannot be the right theory of either epistemic or zetetic norms. In some cases, it’s epistemically and zetetically best to prefer beliefs or acts that predictably generate less accurate total belief states than available alternatives. Intervening to mold the world to our beliefs isn’t zetetically required. And such interventions are at least sometimes zetetically prohibited. If we want to understand the relation between truth and the norms of inquiry, we had better find out why.

References


Thorstad, David (manuscript a). “Inquiry Under Bounds.”

Thorstad, David (manuscript b). “Norms of Inquiry.”