Epistemology vs. Non-Causal Realism

Jared Warren

Abstract

This paper formulates a general epistemological argument against what I call non-causal realism, generalizing domain specific arguments by Benacerraf, Field, and others. First I lay out the background to the argument, making a number of distinctions that are often missed in discussions of epistemological arguments against realism. Then I define the target of the argument — non-causal realism — and argue that any non-causal realist theory, no matter the subject matter, cannot be given a reasonable epistemology and so should be rejected. Finally I discuss and respond to several possible responses to the argument on behalf of the realist. In addition to clearing up and avoiding numerous misunderstandings of epistemological arguments of this kind that are quite common in the literature, this paper aims to present and endorse the first fully rigorous and perfectly general epistemological argument against realism.

Keywords: Epistemological arguments against Realism, Reliability challenge, Realism, Epistemology

Epistemological arguments against this or that brand of realism are nearly as old as the realist positions they attack. In barest outline, these arguments claim that if realism about $X$ is true then we can only be reliable about $X$ by getting lucky. We're left with a choice: reject $X$-realism, reject $X$-reliability, or accept cosmic luck. In recent years, epistemological arguments against realism have appeared in disparate areas of philosophy; but not only do the domains of these arguments vary, so too do the targeted features of realism. Some epistemological arguments against realism target the causal inertness of the objects posited by the realist; others target the mind independence of the facts posited by the realist. Variations proliferate.

Yet the epistemological problems faced by realists seem to be identical in areas as diverse as mathematics, metaethics, and logic; domain specific formulations of epistemological arguments don’t account for this. Hardcore versions
of realism in non-physical domains face exactly the same epistemological challenges, so it is important to get clear on which features of realism generate epistemological problems. This paper attempts to identify the proper structure of the general epistemological argument against realism. My version of the argument is most influenced by the literature stemming from Paul Benacerraf’s famous challenge for mathematical platonism and its reformulation by Hartry Field. But I don’t think that a fully rigorous and perfectly general formulation has hitherto been provided. Here I attempt to provide such a formulation and use it to argue that what I call non-causal realism must be rejected no matter the domain.

1 Reliability

Let $X$-discourse be some branch of discourse, e.g., mathematical discourse, modal discourse, logical discourse, ethical discourse, etc. A sentence is truth-apt just in case it is capable of truth or falsity; Cognitivism about $X$ holds that typical $X$-sentences are truth-apt. When cognitivism holds of $X$ and we have attitudes of acceptance and rejection toward the contents of $X$-sentences, then we need to give some epistemology for $X$: a theory detailing how the attitude-forming mechanisms responsible for our $X$-attitudes operate. By our $X$-attitude forming mechanisms I mean whatever cognitive mechanisms are responsible for the formation of our $X$-attitudes. We can now characterize reliability about $X$ as follows:

Reliability: we are $X$-reliable just in case: (1) The vast majority of $X$-sentences that we accept are true; (2) The vast majority of $X$-sentences that we reject are false; and (3) We accept/reject a significant number of $X$-sentences.

The first two conditions together say that our $X$-attitudes are generally correct and the third says that we have a significant number of $X$-attitudes. At least six points of clarification about this principle need to be made before proceeding:

(i) The notion of reliability used here is quite weak; it involves only an actual correlation between our attitudes and the $X$-truths rather than a counterfactual

---

correlation.

(ii) The “vast majority” quantifier allows that we can make major and important mistakes even about domains in which we’re reliable. This is what we want: we can be generally reliable about a subject without being perfectly reliable about that subject.

(iii) To say that “we” are reliable is to attribute reliability to our community as a whole, but in many cases average members of the community don’t accept or reject any X-sentences. For example: the average person doesn’t understand and so doesn’t have any attitude toward sentences of complex analysis. Still, we can say that our community is reliable about complex analysis because those who have attitudes toward sentences of complex analysis are generally reliable and those without attitudes about complex analysis defer to those with such attitudes.²

The use of “we” also assumes that there is agreement about X among those with X-attitudes. This is an idealization that is tantamount to assuming that there aren’t numerically large outlier groups with X-beliefs. This isn’t always the case, but we can finesse things by noting that some of those with X-beliefs as experts—people that ordinary speakers defer to concerning X. This fends off preposterous results, e.g., we wouldn’t want it to be the case that we, as a community, are unreliable about the age of the Earth if young Earth creationists outnumber geologists.

(iv) This principle doesn’t assume anything substantive about the nature of truth or the truth predicate.

(v) Talking about accepting or rejecting an X-sentence is really a shorthand for talk of accepting or rejecting the content of some X-sentence on a particular occasion of its use. The rejection of a sentence has often analyzed as the acceptance of that sentence’s negation. If this is accepted, then we only need one attitude of belief (acceptance) rather than two (acceptance and rejection). I’m not assuming this account of rejection here because I want the argument I develop to apply quite generally and some proponents of alternative logics think that rejecting a sentence and accepting its negation come apart (e.g., because they think you can’t both accept and reject p but you can accept both p and ¬p).³

²The role of experts and social coordination in content and belief-formation has been widely recognized since Putnam (1975) and Burge (1979).
³See, e.g., the discussion in Priest (2006); Ripley (2011) provides an overview of theories on rejection and its relationship to negation.
(vi) There is some indeterminacy about what counts as a “vast majority” of sentences we accept or reject; or about what counts as a “significant number” of X-sentences. In addition there might be some slack left concerning which X-sentences are really and truly accepted/rejected (e.g., if we accept X-sentence $p$, do we thereby accept $p \land p$, $p \land p \land p$, etc.?). Pursuing such issues would be a distraction here.

2 What is a Reasonable Epistemology?

Not all epistemologies are created equal. One constraint on good epistemologies is that they be reasonable. My terminology here follows a line from Benacerraf, where he says that mathematical realism clashes with:

...the concern that the account of mathematical truth mesh with a reasonable epistemology.\(^4\)

I will define a reasonable epistemology for subject $X$ as one that entails that our $X$-attitudes are both sensitive and safe in a particular sense that I will explain. “Sensitivity” is based on a condition on knowledge in Robert Nozick’s work and “safety” comes from a condition first devised by Ernest Sosa and subsequently discussed and endorsed by Duncan Pritchard and Timothy Williamson among others.\(^5\) Rather than using these notions as standardly formulated, I’m going to alter and generalize them so that they apply not to attitudes toward particular sentences but rather to attitudes toward types of sentences (e.g., mathematical, logical, ethical, modal, metaphysical, etc.):

**Sensitivity**: If, in nearby scenarios, the $X$-truths had been different, then our $X$-attitudes in the counterfactual scenario would be, by and large, correct

**Safety**: If, in nearby scenarios, our $X$-attitudes had been different, then our $X$-attitudes in the counterfactual scenario would be, by and large, correct\(^6\)

Acceptance is correct only for true sentences, rejection only for false sentences. I’m going to make five points of clarification about these principles before putting them to use.

\(^4\)From the first page of Benacerraf (1973).
\(^6\)Both of these conditions could have a clause added to rule out systematically interfering factors, but I’ll ignore this below.
(i) A scenario only counts as “nearby” if, in the relevant scenario, we use the same cognitive mechanisms for forming \(X\)-attitudes that we use in the actual world. This is because we are concerned only with the reliability and functioning of our actual cognitive mechanisms; scenarios where we form mathematical beliefs not with calculations and proofs but instead with coin flips or random guesses are irrelevant to the epistemological standing of our actual mathematical beliefs.

(ii) Even taking (i) into account, there will often be some indeterminacy about which scenarios are “nearby”, but this is standard when dealing with counterfactuals. What matters is that the actual scenario isn’t swamped with possibilities of error (scenarios where our \(X\)-attitudes are incorrect). The import of the “nearby” is that we don’t want to have to change too much in order to generate possibilities of error, e.g., there is natural difference between making sociological changes and biological changes just as there is a difference between making biological changes and physical changes. If we can only generate a possibility of error by tinkering with the laws of physics, that scenario isn’t as nearby as one in which only some sociological facts have been changed.

(iii) I’m not making any substantive assumptions about the kind of modality involved in the counterfactuals in the Sensitivity and Safety conditions. For all I’ve said here it could be conceptual or metaphysical or something else altogether, but the usual assumption is that standard counterfactuals involve metaphysical modality.

(iv) The “by and large” clause is vague; how many possibilities of error is too many? Since we aren’t literally counting scenarios, it’s impossible to give a precise, quantitative answer. The rough idea is that if it is easy to come up with possibilities of error similar to the actual world then we have a problem.\(^7\)

(v) By talking about “scenarios” I don’t mean anything fancy or technical. The talk is just a surrogate for the overused and perhaps misleading talk of “worlds” that is standard when using counterfactuals and modal operators. Here and below I will talk indifferently of “scenarios” and “worlds”.

\(^7\)Some sensitivity, safety, and anti-luck conditions in epistemology are much stricter, e.g., Nozick (1981), pages 172-196, and Heller (1999)’s account of epistemic luck.
3 Why Do We Need a Reasonable Epistemology?

The demand for a reasonable epistemology amounts to the demand that our epistemologies be both sensitive and safe when we’re reliable. But while some versions of these generalized conditions are tacit in the literature on epistemological arguments, as far as I know, the only writer in the literature on this topics who has explicitly introduced something like both of these constraints is Justin Clarke-Doane who, in recent work, endorses the:

SS Principle: Information, E, cannot undermine our beliefs of a kind, F, without calling into question our ability to show that our F-beliefs are both safe and sensitive.\(^8\)

But in demanding a reasonable epistemology, I am arguing for something like the converse of Clarke-Doane’s principle: if our \(X\)-beliefs are not both sensitive and safe then they are undermined.

The sensitivity portion of this seems widely accepted in the literature on epistemological arguments, but it is also clear that there is something wrong with an epistemology that fails the safety condition. Imagine we manage to get our mathematical beliefs correct simply by throwing darts at board with mathematical claims written on it. In this case, our mathematical beliefs would perhaps be sensitive, since there are no scenarios at all where the necessarily true mathematical truths differ, there are, a fortiori, no nearby scenarios where they differ.\(^9\) So in the imagined case we would be both sensitive and reliable about mathematics, yet surely there is obviously something epistemically wrong with our mathematical beliefs in this scenario. I think the problem with these beliefs is simple: in nearby scenarios where the darts we threw hit slightly different parts of the board, our mathematical beliefs differ and we fall into error, i.e., these beliefs aren’t safe.

There are different ways to motivate the requirement that our \(X\)-beliefs need to be both sensitive and safe. One is to endorse sensitivity and safety as conditions on knowledge. Then, if we can show that our theory of \(X\) shows that

\(^8\)Quoted from Clarke-Doane (forthcoming). Clarke-Doane then argues that our mathematical beliefs are both sensitive and safe, so they are not undermined (this represents a change of mind from Clarke-Doane (2012). This isn’t strictly relevant though, for proponents of anti-realist arguments aren’t arguing that our mathematical beliefs are not both sensitive and safe, but rather that if mathematical realism is true, then our mathematical beliefs are not both sensitive and safe. That epistemological arguments try to undermine certain philosophical theories rather than certain subjects is missed over and over again in the literature, e.g., see Burgess & Rosen (1997) and White (2010).

\(^9\)Cf. Linnebo (2006)’s “Lucky fool”.
our beliefs are not sensitive or not safe, we will have shown that our theory of
X entails that we can’t have X-knowledge. Another strategy is to argue that
our X-reliability calls out for some kind of explanation since it is a striking
alignment between our beliefs about X and the X-truths and this alignment
can’t plausibly be taken as brute. It could then be argued that the only way to
explain our X-reliability is to show that both the sensitivity and safety conditions
obtain. According to this approach, sensitivity and safety are motivated as
requirements on explanations.

I am not going to endorse either of these strategies here, since we can rest
with the fact that our X-reliability is extremely unlikely if our X-attitudes aren’t
both sensitive and safe. If the conjunction of X-reliability and non-causal re-
alism about X is unlikely in this manner, then X-reliability can be used to
put pressure on the acceptance of non-causal realism about X. When our X-
attitudes aren’t both sensitive and safe, it’s fair to say that our X-reliability is
simply a matter of luck. Without sensitivity and safety, the actual world, in
which we are reliable, is swamped by possibilities of X-error. There is natural
theoretical pressure against accepting this kind of cosmic luck. While it might,
at first glance, sound a bit tendentious, the anti-luck principle invoked here is
highly intuitive, plausible, and nearly undeniable when properly understood.
Let’s say that we are X-lucky if we are X-reliable without a reasonable episte-
ology for X being possible:

No Luck: for every subject X, we are not X-lucky, i.e., for any
subject X either we aren’t X-reliable or a reasonable epistemology
for X is possible

It is crucial that this principle not be misunderstood. In particular, the three
following points of clarification about No Luck are key.

(i) Sometimes people get lucky, like when Lucky Larry wins the lottery.
Wouldn’t reasoning analogous to what led us to the No Luck principle also
lead us to reject that Lucky Larry won the lottery? Emphatically no. Of course
not, but since this type of objection is often raised, it’s important to understand
why. The key is that in Lucky Larry-style cases we have a ton of independent
evidence that the unlikely event happened, by contrast, in the theoretical cases
we are considering, we have no such independent evidence for the conjunction
of non-causal realism about X and X-reliability.\footnote{Setiya (2012)’s discussion of epistemological arguments misses this crucial point on pages 68-69.} It isn’t like we can peek be-
hind the cosmic curtain and independently assess the $X$-truths as they are in and of themselves, as viewed by God. Failure to grasp this clarification will lead to spurious objection to No Luck.

(ii) $X$-reliability leads to only to the demand that a reasonable epistemology for $X$ is possible. This is very different from demanding that we already possess a reasonable epistemology for $X$. There isn’t any problem with our not currently possessing a reasonable epistemology for $X$. The perceptual beliefs of cavemen weren’t somehow undermined because they couldn’t provide a reasonable epistemology for those beliefs given their meager resources at the time.

(iii) There is a huge literature on epistemic luck that I am sidestepping.\footnote{For overviews see Pritchard (2005) and Engel Jr. (2014).} Historically there is agreement that certain kinds of epistemic luck undermine our beliefs (usually by undermining our knowledge) while other kinds don’t. Kinds that don’t include the luck of merely existing, of having the right cognitive mechanisms for forming the relevant beliefs, or the luck of experiencing rare events (like in (i)).\footnote{The first explicit recognition of these non-undermining forms of luck is in Unger (1968).} But the luck needed for reliability when sensitivity or safety fails involve scenarios where our beliefs are of the same kind, produced by the same cognitive mechanisms and these certainly matter for assessing the worth of these cognitive mechanisms for the particular type of belief being considered. So we don’t need to accept any controversial luck-based analysis of knowledge or accept some unduly strong construal of “epistemic luck” in order to accept and work with the No Luck principle. Again, failure to grasp this clarification leads to confusing the plausible and straightforward No Luck principle with a much more dubious claim.

With these clarifications in place, we can now put our anti-luck principle to work. First note that any epistemology for subject $X$ will have to build in some philosophical theory $T$ about $X$. $T$ will tell us things like whether or not the $X$-truths are mind dependent or independent, whether they can be accessed causally, etc. (see the following section for more on this). Now consider the following schematic argument:

1. If a reasonable epistemology for $X$ is not possible according to theory $T$, then either we are not $X$-reliable, $T$ is false, or we are $X$-lucky (this premise follows from what has been said above)

2. We are not $X$-lucky (No Luck)
3. So: if a reasonable epistemology for $X$ is not possible according to theory $T$, then either we are not $X$-reliable or $T$ is false \((1, 2)\)

Since the conclusion of this argument summarizes the discussion of adopting a sensitivity and safety constraint, let’s give it an appropriate name:

**SS Constraint**: if a reasonable epistemology for $X$ is not possible according to theory $T$, then either we are not $X$-reliable or $T$ is false

This principle will play a crucial role in the general epistemological argument against non-causal realism, but before proceeding to that argument, we must get clear on its target.

4. **The Target**

The target of my argument is what I call “non-causal realism”. There are several formulations of “realism” in the philosophical literature and I have no desire to engage in proprietary squabbles over terminology, so my uses of the terms *realism* and *non-causal realism* are stipulative. I feel the need to stress this at length, since despite my protests, several philosophers have responded to my arguments here by saying things like “but should this other theory count as “realism”, even though it doesn’t according to your definition?” Perhaps it should. The point here is not to characterize the essence of the term “realism” for all time, but only to cleanly segment theories into those challenged by my epistemological argument.

With these pleas out of the way, let us start getting clear on the target: $X$-realism must embrace **Cognitivism** about $X$. In addition we don’t want an **error theory** about $X$—the idea that all atomic $X$-sentences are trivially false—to count as a version of $X$-realism:\(^{13}\)

**Non Vacuity**: There are some non-trivial $X$-truths

Next, realism requires that the $X$-truths be **mind independent**:

**Mind Independence**: The $X$-truths are mind-independent

The required notion of **mind independence** is difficult to characterize precisely. It’s natural to try for a modal characterization, e.g., by saying that the $X$-truths

\(^{13}\)For error theory about ethics, see Mackie (1977); for error theory about mathematics, see Field (1989).
are mind independent just in case they don’t counterfactually depend upon any human mental states, but there are various semantic tricks that can be used to deliver these counterfactuals even within a philosophical theory of X that clearly makes the X-truths mind dependent. Because of this, it is perhaps better to define mind independence using some general hyperintensional dependence relation instead.\(^\text{14}\) But here I’m just going to ignore rigidifying tricks and assume that the X-truths are mind independent just in case they don’t supervene upon the mental truths, i.e., just in case differing the X-truths need not result in any difference to the mental truths.

The last requirement on X-realism is that what I call easy pluralism about X is rejected. This has two components. First, X-pluralism is the idea that there are multiple equally correct theories of X. So a logical pluralist might maintain that both classical logic and intuitionistic logic are correct theories of logic, and an ethical pluralist might maintain a theory entailing that killing is always wrong and a theory entailing that killing is sometimes okay can be equally correct. To make this idea clear we need to know what makes something a “theory of X”. So, e.g., what makes both classical logic and intuitionistic logic count as theories of logic even if they don’t overlap at all in the symbols they use or are framed in languages with wildly different grammatical structures? I think the best way to approach this question is localize the discussion a bit further, by looking at various logical notions, e.g., negation, conjunction, the conditional, etc. and to say that any theory of these notions counts as a theory of logic. This replaces the problem of saying when two theories both count as theories of X with the problem of saying when two notions are both \(\alpha\)-like notions when \(\alpha\) is a notion characteristic of X. Clearly both classical negation and intuitionistic negation are, e.g., negation-like notions (and similarly for the other logical notions), but how are we to make sense of this?

The natural approach to this problem is to look at the wide inferential role that the symbol for negation (“\(\neg\)”) plays in both the classical theory and the intuitionistic theory. These symbols don’t obey the same (structurally individuated) laws in these theories, and so they don’t have an identical inferential role in both theories, e.g., in classical logic \(\neg\neg\phi\) will be fully equivalent to \(\phi\), but not in intuitionistic logic. Still, the inferential role played by the negation symbol in the classical theory is relevantly similar to the role played by the negation sym-

\(^\text{14}\)See Fine (2001) and Jenkins (2005).
bol in the intuitionistic theory, so that we may say that the notions are semantic counterparts.\textsuperscript{15} It need not be the case that each theory has a counterpart for every notion expressible in every other theory; and it need not be the case that each theory has only one counterpart for a given notion in another theory. To be semantic counterparts is to enjoy similar wide inferential roles, and like all similarity relations, ours is subject to the difficulties of how to weigh and tally the various similarities and dissimilarities in particular respects. The right way to tally will depend on what is relevant and salient in a given context, so what counts as a semantic counterpart for some purposes and in some situations may not so count for other purposes and in other situations.

A theory $T$ in some language is a theory of, e.g., logic, if $T$ includes semantic counterparts of our various logical notions and two theories of logic are apparently incompatible if they disagree with each other when semantic counterparts are translated into each other. So we can easily see that classical logic and intuitionistic logic are apparently incompatible theories of logic (an identical story can be told for any subject $X$). It’s important that the relevant notion of incompatibility not be misunderstood. We don’t want it to be the case that, e.g., set theorists who deny the axiom of choice aren’t allowed to say that there are sub-universes in which the axiom of choice is true. In this case, we are only concerned with their unrestricted set theory, not with any restriction of one of their theories. This condition isn’t perfectly precise, but there is no mystery about how to apply it in practice.

Easy pluralism is the idea that had we adopted an alternative theory of $X$, then that alternative theory would have been correct (i.e., true) in our alternative language:

\textbf{Easy Pluralism}: in nearby scenarios where we adopt an alternative theory of $X$ incompatible with our actual theory of $X$, that alternative theory is true in the language spoken in that nearby scenario.

There are many reasons that someone could endorse a version of easy pluralism about $X$. For instance: if someone was an $X$-conventionalist and took the $X$-truths to be rules of language or analytic truths, then adopting an alternative theory of $X$ would be tantamount to endorsing different linguistic rules or analyticities and thus having different $X$-truths.\textsuperscript{16} Or one might endorse

\textsuperscript{15}I take the term “counterpart” from its use in counterpart theory; see Lewis (1968).
\textsuperscript{16}See Ayer (1946), Carnap (1934), and Giannoni (1971) for versions of conventionalism about logic and mathematics.
a charity-heavy meta-semantics, where charity to use involve attributing truth
to our counterfactual selves’ corely held beliefs, including their alternative X-
axioms.\footnote{See Davidson (1984), Hirsch (2011), Lewis (1974), and Quine (1970) for more on charity
principles of this kind.} But X-realists cannot endorse easy pluralism; it is essential to realism
of the kind targeted by my argument that there be a real risk of getting the X-
truths wrong, and easy pluralism forecloses upon this possibility.\footnote{I think it is plausible that we may be able to let easy pluralism do the work of mind indepen-
dence in the definition of realism; in that case though we would arguably need to add a condition
that the physical truths don’t subvene upon the X-truths, to the causal inertness condition below. I
won’t explore this possibility any further here.} N.B., easy pluralism is indefinitely easy: a view can be a version of easy pluralism and still think that we can make mistakes about X of various kinds.\footnote{Though I won’t discuss this here, I think the best way to define pluralism in full detail is to use
the notion of a semantic counterpart, see my (2014) and (2015) for discussion.}

We can now define X-realism:

**Realism** : A theory of X, T, is a version of X-realism just in case T
endorses Cognitivism, Non Vacuity, Mind Independence, and
rejects Easy Pluralism

Realism can be endorsed without epistemological mystery in domains where
we can causally access the relevant truths, e.g., ordinary objects like tables and
chairs are things that almost all of us are realists about. This is unproblematic
because we can sketch, at least in broad strokes, a reasonable epistemology for
our beliefs about tables and chairs involving perception and our causal interac-
tions with the world. I will argue that realism (so defined) only generates an
epistemological problem when some kind of acausality condition is assumed:

**Causal Inertness** : the X-truths are causally inert

Although an intuitive understanding of this principle would suffice for my pur-
poses, it’s worth trying to clarify it a bit. Benacerraf’s original epistemological
argument assumed a now widely rejected causal theory of knowledge, but de-
spite the failure of this theory, mathematical realists and many other sorts of
realists endorse some kind of causal inertness condition.\footnote{See Goldman (1967) for the introduction of a causal theory of knowledge; and see Carrier
(1993) for a more recent endorsement and development of a causal theory.}

What is it for X-truths to be causally inert? Minimally, causally inert things
don’t enter into causal interactions either as causes or as effects, but that’s much
too weak as a characterization, since many philosophers believe that only events
enter directly into causal relations. If this were right, then the just-given characterization would entail that people, tables, baseballs, etc. are causally inert, and that is clearly the wrong intuitive result.

It’s difficult to correctly analyze the notion of causation, but it’s generally accepted that fundamental particles paradigmatically enter into causal intercourse with each other and that non-fundamental physical entities have, at least, derivative causal powers inherited from the causal powers of their fundamental parts. Here causation is meant as a covering notion for all kinds of difference-making physical interactions. Even those theorists who, strictly and literally, reject causation should be able to accept my claims about “causation” as long as they accept difference making physical interactions (and it’s difficult to see how anyone could completely reject such interactions). This gives us the leeway to use a bazooka rather than a scalpel when characterizing causal inertness: the X-truths are causally inert when the physical truths don’t supervene upon them, i.e., when there can be a difference in the physical truths without there being any difference in the X-truths.21

Now, at long last, we can define our target:

Non Causal Realism: a theory of X, T, is a version of non-causal realism about X just in case T is a version of Realism and T endorses Causal Inertness about X

It is important to note that, according to my usage, non-causal realism doesn’t require any form of ontological realism about X; this means that it can be endorsed in domains that don’t make distinctive ontological claims (e.g., logic, ethics). This is a crucial feature that will, unlike the arguments of Benacerraf and Field, allow for epistemological arguments to be mounted against non-causal realism in domains like ethics, logic, and modality. It’s also one reason endorsing Causal Inertness about X isn’t tantamount to endorsing a causal theory of knowledge for X.

Despite the fact that the epistemological argument I give is explicitly targeting only non-causal domains (in the relevant sense) I am still occasionally faced with objections concerning whether my argument applies to, e.g., material objects as they existed in the distant past. But obviously, facts concerning material objects are not causally inert. As the next section will demonstrate: the causal

---

21It’s perhaps worth noting that I’m not taking sides in the dispute between Humeans and non-Humeans here: I’m simply building causal structure into the notion of physical truths. It may well be that the physical truths will include truths concerning natural laws in addition to the truths that make up the Humean mosaic.
inertness condition is required if my argument is to go through. In any event, the type of epistemological challenge that I’m posing is only for non-causal realism, which, as we’ve seen in this section, involves rejecting easy pluralism about X in addition to accepting the causal inertness and mind independence of the X-truths.

Non-causal realism isn’t a strawman; some serious philosophical theories are versions of non-causal realism. The non-naturalist ethical realism endorsed by David Enoch, Thomas Nagel, Derek Parfit and others is a version of non-causal realism, as is the traditional brand of mathematical platonism endorsed by Kurt Gödel, Hugh Woodin, and others.\textsuperscript{22} And some recent metaphysicians, including Kit Fine, David Lewis, and Theodore Sider, seem to endorse non-causal realism about mathematics, logic, ethics, and many other domains as well.\textsuperscript{23} I don’t want to examine or discuss any of these particular theories, only to stress that various versions of non-causal realism are seen as live options in contemporary philosophy.

5 The Argument

The argument against non-causal realism needs one further premise expressing a very weak form of naturalism about human cognition:

\textbf{Naturalism}: all of our cognitive mechanisms are causal mechanisms, i.e., attitudes change only in response to causal changes of some kind, no matter the attitude or subject under consideration

\textit{N.B.}, this is an extremely weak form of naturalism, it is naturalism about cognition and doesn’t in any way assume that all objects we talk about can be identified with some part of the physical world. Nor does it even assume that all of the outputs of brain activity are causal, i.e., while many philosophers have rejected this form of naturalism about consciousness, very few have rejected the kind of naturalism about cognition summed up in this principle.\textsuperscript{24} \textbf{Naturalism} is so obvious that it’s difficult to effectively argue for it. Anyone who rejects this principle thinks that our attitudes about some subject matter can change without there being any causal change, e.g., we can believe some claim \( p \) at time \( t \)

\textsuperscript{22}\textsuperscript{22}See Enoch (2011), Nagel (1986), Parfit (2011a) and (2011b), Gödel (1964), and Woodin (2004).
\textsuperscript{23}\textsuperscript{23}See Fine (2001), Lewis (1984), and, especially, Sider (2011).
\textsuperscript{24}\textsuperscript{24}Though see the closing section of Boghossian (forthcoming).
and then not believe \( p \) at a later time \( t+ \) even though the entire physical world is identical at \( t \) and \( t+ \). This is, to say the least, mysterious. In any case, this type of cognitive naturalism is a presupposition of our modern scientific view of the human mind and my argument assumes it.

With this further premise, we can show that a reasonable (sensitive and safe) epistemology for \( X \) cannot be given if non-causal realism about \( X \), for any \( X \), is accepted. I will give arguments that both sensitivity and safety fail, and I think both of these arguments are correct, but recall that all that we need, strictly speaking, is for at least one of these arguments to succeed.\(^{25}\)

(i) **Sensitivity.** Assuming that the \( X \)-truths are different in some nearby scenarios, by the conditions of mind independence and causal inertness, the nearby scenarios in which the \( X \)-truths are different are scenarios in which all of the physical and mental truths remain the same, i.e., these nearby scenarios are scenarios that are physically and mentally identical to the actual world. But then, by cognitive naturalism, our beliefs in these counterfactual scenarios must be exactly the same as they are in the actual world. So, since the truths are different and our beliefs the same, we fall into massive and widespread error about \( X \). So assuming that there are nearby scenarios where the \( X \)-truths differ, by causal inertness and mind independence, our \( X \)-attitudes won’t be sensitive. It seems obvious that there are such scenarios, and here I will assume that there are, postponing discussion of this to section 6 below.

(ii) **Safety.** The failure of the safety constraint on non-causal realism follows directly from the rejection of easy pluralism clause together with the premise that there are nearby scenarios in which our \( X \)-attitudes are different. The anti-pluralism condition holds directly that if there are such situations, they are situations in which our counterfactual selves fall into error (i.e., have many false \( X \)-beliefs). This is virtually what the anti-pluralism clause says, viz., there is only one true \( X \)-theory, so provided that we have a true \( X \)-theory in the actual world, in any counterfactual scenario where we have a different \( X \)-theory we have a false \( X \)-theory.

Establishing that there are nearby scenarios where our \( X \)-beliefs are different involves assumptions about the subject under consideration, but in many cases it is relatively straightforward. For example: for the case of logic, there are many different logics that have been embraced by philosophers: classical logic, intuitionism, paraconsistent logics, quantum logics, etc. It is easy to imag-

\(^{25}\) The possibility of working with just the safety condition is mooted below, in section 6, and I am sympathetic, but will officially stick with the stronger double condition.
ine an alternative scenario where we reason according to one of these alternative logics. According to the realist, in such scenarios, we fall into logical error. And in the case of ethics there are actual differences in ethical codes between human societies past and present. So if we are correct that slavery is morally wrong, and we aren’t pluralists, then plantation owners in the nineteenth century American south and citizens in ancient Athens were incorrect in assuming that slavery was morally permissible. If there can be actual moral differences, then clearly there are possible moral differences. In worlds where the South prevailed in the American civil war and slavery is widely sanctioned and accepted, we—according to moral realists—fall into widespread error.

It’s obvious that in all of the cases that concern us, realists must admit nearby scenarios where we fall into error. They cannot deny this, since claiming that our X-beliefs could not have been different is dangerously close to claiming that our X-beliefs are analytic or conceptually true, and accepting this amounts to accepting a version of easy pluralism about X, which is incompatible with X-realism. So the non-causal realist must admit that no safe epistemology for X can be provided, according to their view.

With this established, we can finally formulate the master argument against non-causal realism in domains of X-reliability (in schematic form):

1. We are X-reliable (premise)
2. If a reasonable epistemology for X is not possible according to theory T, then either we are not X-reliable or T is false (SS Constraint)
3. A reasonable epistemology isn’t possible for X according to non-causal realism about X (just argued for)
4. So: non-causal realism about X is false (1, 2, 3)

A key point about this argument, one that is very widely misunderstood in discussions of arguments like this, is that the conclusion of the argument is that a particular philosophical theory of X is false. Not that our X-beliefs themselves are somehow undermined. Epistemological arguments never undermine our mathematical or ethical beliefs, rather, they undermine them on the assumption of a particular type of philosophical theory. As I noted in passing above, this is often missed in the literature. This key point is what shows that these epistemological challenges aren’t skeptical in a pejorative sense.\(^{26}\) This confusion

\(^{26}\)More evidence that the challenge is not skeptical is provided by pointing out that causal epistemologies can be given for ordinary material objects.
has probably done more to pollute philosophical discussions of epistemological arguments against realist than any other; I think that one of the key contributions of my more general and more explicit treatment is that it makes it much, much easier to identify and avoid this and related confusions when discussing epistemological arguments.

6 Realist Rejoinder 1: Necessity

One extremely well-known response to anti-realist epistemological arguments in the literature I call the necessity response; it was first given by David Lewis in response to Benacerraf. Lewis claimed that since mathematical truths are metaphysically necessary, there are no possible worlds in which the mathematical truths are other than they are, and so our mathematical beliefs are trivially sensitive. This is because—according to the standard semantics for counterfactual conditionals—any counterfactual with a necessarily false antecedent is trivially true.

The necessity response has been widely criticized in the literature on the Benacerraf/Field problem, but for a variety of disparate reasons. There is no agreement about why it fails, only that it does fail. Here though we can see that even if we grant everything Lewis says, the necessity response will only guarantee that our \( X \)-beliefs are sensitive; it does nothing at all to show that they are safe. And no similar move could work to guarantee safety, since worlds in which we have differing \( X \)-beliefs, no matter the \( X \) in question, are always not just metaphysically, but physically possible.

So even granting the success of the necessity response at guaranteeing sensitivity (something that is controversial), it doesn’t succeed at answering my argument on behalf of the realist. And the Lewis response again nicely illustrates how little we have achieved if we have sensitivity alone without safety. But granting, for the sake of argument, the success of the necessity response at achieving sensitivity, the realist’s task of answering the argument is made much

---

28 I’m extrapolating a bit from Lewis here—it sometimes sounds like he’s saying that sensitivity isn’t needed in domains of necessary truth, but at others it seems like he’s saying that sensitivity is trivial in such domains. The differences between these readings aren’t important here.
29 The standard semantics are all based on one or both of the slightly different accounts in Stalnaker (1968) and Lewis (1973).
30 Perhaps the most popular response is that pace standard formal accounts of counterfactuals, some natural language counterfactuals with necessarily false antecedents aren’t trivially true or incoherent. See Field (1989), pages 233-242.
easier—now they need only provide a safe epistemology for X.\textsuperscript{31}

7 Realist Rejoinder 2: Evolution

One strategy for supplementing the necessity response’s deliverance of sensitivity is to use an evolutionary account of our mechanisms for forming X-beliefs to show that those beliefs are safe. Recently several non-causal realists have attempted to answer the epistemological challenge by appealing to evolutionary considerations.\textsuperscript{32} In outline, an evolutionary account of our X-reliability runs as follows: the cognitive mechanisms responsible for our X-beliefs were selected for their ability to form true X-beliefs, i.e., having a cognitive mechanism that is reliable about X conferred some kind of evolutionary advantage upon our ancestors and this evolutionary story explains why we tend to have true X-beliefs.

But even if the relevant cognitive mechanisms were selected for, it doesn’t follow that they were selected for their ability to produce true beliefs. To illustrate this, let’s consider a distinction made by Hartry Field between:

1. Selection pressure for acceptance of a particular X-theory T, which is in fact correct
2. Selection pressure for acceptance of the correct X-theory, whichever one that happens to be correct\textsuperscript{33}

(1) isn’t enough to provide a Darwinian vindication of non-causal realism, for that, the stronger claim (2) is needed. If (2) were supplied, we would be able

\textsuperscript{31}The necessity response can be supplemented by a plenitude theory about X, i.e., by the claim that roughly any X-theory that we had chosen to accept would have been correct because the X-facts are so abundant and meta-semantics so charity heavy. The only example of this in the literature that I am aware of is Mark Balaguer’s \textit{Plenitudinous or Full-Blooded Platonism} in the philosophy of mathematics; see Balaguer (1995) and (1998). Even though I think it succeeds in showing that our mathematical beliefs are safe, Balaguer’s view isn’t relevant here, since it endorses easy mathematical pluralism and thus isn’t a version of non-causal realism. An anonymous referee objected to this claim, saying that Balaguer’s views “intuitively” seems like it should be a version of non-causal realism. I am totally puzzled by this, since Balaguer’s view is explicitly a version of pluralism and thus not a version of non-causal realism. I think that the referee’s objection was just that Balaguer’s view seems like it should be a version of “realism”. Of course, call it whatever you like, and it is a version of (e.g.) \textit{ontological} realism about mathematics, the point is only that Balaguer’s view is a version of pluralism and so doesn’t fall into my target area. To think that easy pluralism requires indefinitely easy pluralism is to miss this point; there is a real though somewhat hazy distinction between pluralist views like Balaguer and non-pluralist views (see also Field (2005) and Schechter (2010) for agreement with this point).

\textsuperscript{32}See Schechter (forthcoming) and Enoch (2010); by contrast, Street (2006) uses evolutionary considerations to argue against normative realism.

\textsuperscript{33}Field (2005), pages 80-81; Field is just discussing logic, I have generalized.
to show that there were no nearby scenarios where we had distinct $X$-beliefs and thereby fell into error. This would only get us safety, but since I’ve granted that the necessity response can get us sensitivity, together these would provide a successful response to the argument.

Unfortunately for the non-causal realist, no evolutionary story vindicating (2) is plausible in non-causal domains. If a slightly different evolutionary path would have resulted in our having different $X$-beliefs while having the same cognitive mechanisms, then a slightly different evolutionary path would have resulted in us being wrong about $X$ according to non-causal realism. So an evolutionary account of our $X$-reliability suitable to the defense of non-causal realism must show that it wasn’t possible for us to have different $X$-beliefs other than those that we actually have, and this cannot be done. The general problem is that selection forces are blind to distinctions between distinct realist theories. So if $T$ is our theory of $X$, and $T^\ast$ is a slightly different theory of $X$ that would bestow upon us the same survival and reproductive advantages as $T$, in any alternative evolutionary history where we accept $T^\ast$ rather than $T$ we will fall into massive error. It will always be the case that many different theories are such that their acceptance would have had roughly or exactly the same survival benefits for our ancestors (accepting modus ponens and accepting modus ponens only on Earth, or for conceptually simple beliefs, etc. would have conferred the same survival and reproductive advantages upon our ancestors).\textsuperscript{34}

Selection forces are too coarse-grained to be of help to the realist here; there are too many ways of varying our social and evolutionary history slightly and coming up with alternative logical, mathematical, or ethical beliefs. So there will always be too many possibilities of error here for our beliefs to be safe if non-causal realism is accepted.\textsuperscript{35}

\textsuperscript{34}David Enoch (2010) suggests a different evolutionary response to the argument as specifically directed against ethical realism: according to Enoch, survival is good, and this is supposed to make the alignment of our ethical beliefs and the ethical facts less mysterious while not completely dispelling the mystery. I don’t discuss this response directly because I fail to see how it helps the non-causal realist, for roughly the reasons discussed in the main text.

\textsuperscript{35}One response that I won’t discuss in detail is the idea that the argument above keeps bad company: if it worked for logic, math, and ethics, it would work for ordinary objects like tables and chairs as well (this is pushed most forcefully in Clarke-Doane (forthcoming)). I agree with this, but it is irrelevant, since nobody accepts non-causal realism for ordinary objects like tables and chairs, instead we all accept causal realism about tables and chairs. Failure to see this once again arises from confusing the epistemological argument against realism with an argument against reliability in a particular subject.
8 What Now?

I have argued that the epistemological argument against non-causal realism is sound; but this might be rejected while admitting that the argument is valid. If this is done, we're left with the following options for each subject X:

1. Accept that we were lucky in being $X$-reliable
2. Reject our $X$-reliability
3. Reject non-causal realism about $X$

(1) is a complete non-starter in the areas that concern us, for reasons given above. That leaves us with (2) and (3).

Option (2) could be acceptable in cases where we cannot fathom what it would be like if non-causal realism about $X$ were false. The only somewhat plausible example of this that I can come up with concerns the existence and attributes of God. If God doesn't interfere in the world either due to non-existence or by choice then the God-truths will be causally independent of the world in the relevant sense (this would be difficult to maintain if God created the world, but let's ignore this). It's very difficult for me to see what it would even mean to reject realism about God's existence. Accordingly, it seems much more plausible for to reject our reliability in God-discourse by taking the God-truths to simply be beyond our ability to discover. Of course we might get lucky and get some of the God-truths right by mere chance, but as our God-theory becomes more and more elaborate, the pressures against luck start to mount and we are better off rejecting our reliability about God.

In all of the other cases that are of concern to us (mathematics, ethics, logic, modality, metaphysics, consciousness, aesthetics, etc.) I think we should opt for option (3). According to philosophical folklore, non-causal realism is the proverbial person in the street's default view in many areas; but if this is true (and I doubt that it is), then the person in the street has made a mistake. The mistake was in theorizing in these domains by analogy to ordinary objects, which we have causal access to via our causal cognitive powers. It has often been assumed that these subjects concern domains of truths that are just like the truths concerning ordinary objects, save for being causally inaccessible; but this assumption is mistaken and fatal, for it leaves us without an ability to provide a reasonable epistemology in any of these areas. Ergo non-causal realism should
almost always be rejected, no matter the subject matter.\textsuperscript{36}

\textbf{References}


\textsuperscript{36}Thanks to Justin Clarke-Doane, Yu Guo, Sharon Street, Daniel Waxman, and several referees.


