Fast Science

Jacob Stegenga

Draft of January 11, 2024.
Forthcoming, *The British Journal for the Philosophy of Science*

Abstract

If scientists violate principles and practices of routine science to quickly develop interventions against catastrophic threats, they are engaged in what I call fast science. The magnitude, imminence, and plausibility of a threat justify engaging in and acting on fast science. Yet, that justification is incomplete. I defend two principles to assess fast science, which say: fast science should satisfy as much as possible the reliability-enhancing features of routine science, and the fast science developing an intervention against a threat should not depend on the same problematic assumptions as the fast science which estimates the magnitude, imminence, and plausibility of the threat.

1. Science in Supreme Emergencies

The tortoise taught us tenacity triumphs, while the hare's hubris handed him heartbreak. Routine science is a tortoise. Is fast science a hare? A supreme emergency can motivate scientists to cut corners to quickly find a response. The epidemiological modelling at the start of the Covid-19 pandemic is a portentous example: this work was done rapidly, based on empirical assumptions that were soon criticised, promulgated without peer review, presented quickly to policy-makers, and formulated in terms of maximal urgency to elicit specific policies, and the resulting response—lockdown—was one of the most impactful science-informed policies in decades. The putative justification for violating the principles and practices of routine science was obvious: many millions of people could otherwise die. The magnitude, imminence, and plausibility of a supreme emergency warrants fast science. Yet, the hare was too confident in his talents, so he took a nap, and lost the race. My aim is to articulate and defend two principles to keep fast science more like the tortoise than the hare.

Routine science has a range of principles and practices that render it relatively reliable. What exactly those principles are has been a matter of dispute: Popper held that genuinely scientific theories should be falsifiable; Merton called his four principles communism, universalism, disinterestedness, and organised skepticism; Longino emphasises the importance of criticism, including the need for recognised avenues of criticism, shared standards, equality of intellectual authority, and responsiveness to criticism. However one characterises the reliability-enhancing principles of science in such general terms, many routine scientific practices are meant to instantiate those principles. These include long and rigorous traineeships, demanding standards for empirical methods and analysis, competitive allocation of research resources, critical peer review and modification of work in response to review, publication of results, and criticism of existing work. Routine science takes time. The organised skepticism of science entails that scientists usually view their results tentatively. Moreover, most routine science is intended for peers rather than the public or policymakers. Of course many cases of routine science are not ideal, but nevertheless these practices are typical of routine science at its best. Fast science violates some of the reliability-enhancing principles and practices of routine science.

The putative justification to engage in and act on fast science is compelling. If a threat is catastrophic, imminent, and plausible, we ought to quickly learn as much as possible about the threat and possible ways to intervene on it. In response to great threats doing fast science may be better than doing no science, even if the fast science is less reliable than routine science. Drawing on Walzer's notion of 'dirty hands', which holds that political and military leaders can violate routine moral codes in response to a supreme emergency ([1973]), Birch argues that scientific policy guidance can violate routine norms during periods which he calls *in extremis*—times in which the lives of many people are immediately threatened and there exists no prepared plan on how to manage the threat ([2021]). Specifically, Birch argues that while scientists should usually limit themselves to 'normatively light advice' in which scientists offer no specific policy recommendation, when *in extremis* science can provide 'normatively heavy advice' in which scientists do offer specific policy recommendations. Plausibly, the same justification can warrant fast science. Engaging in and acting on fast science is justified, one might say, when science advisors work *in extremis*. While Walzer's claim that supreme emergencies can justify the violation of moral principles might be contentious, that supreme emergencies can justify the violation of scientific principles is much less so. White et al. agree, claiming that normal epistemic standards can be violated during a supreme emergency ([2022], p. 12)].

A supreme emergency is constituted by a 'threat tripod' of magnitude, imminence, and plausibility. The magnitude gives some indication of the scale of the catastrophe were we to not engage in or act on fast science. Imminence gives some indication of the temporal restrictions within which fast science can operate. Walzer's notion of supreme emergency was based on a threat's magnitude and imminence ([2000]), yet the threat tripod is incomplete without plausibility, because threats have varying degrees of plausibility, and, like the estimated magnitude of a threat, a threat's plausibility indicates what things would be like were we to not engage in or act on fast science (if the plausibility of a threat is low, then things would probably be much the same as they are now without fast science, whereas if the plausibility of a threat is high, then things could be very different than they are now without fast science). The plausibility leg of the tripod can be understood as a probability of some minimal degree arising from uncertainty: in the spring of 2020 the emerging pandemic was a plausible catastrophe, but a plague of zombie unicorns was not. The magnitude and plausibility of a threat are standard considerations in risk assessment, and a supreme emergency has the added feature of imminence.

The three legs of the threat tripod must all be tall enough for a threat to be a supreme emergency. If a threat is trifling but imminent and plausible—say, in my fridge tomorrow there will be no milk for tea—then the magnitude leg of the tripod is too short and so the tripod does not stand, there is no supreme emergency, and no fast science is warranted. If a threat is far off in time yet catastrophic and plausible—say, in a few billion years the sun will turn into a red giant and engulf Earth—then the imminence leg of the tripod is too short and so the tripod does not stand, there is no supreme emergency, and no fast science is warranted. If a threat is implausible yet catastrophic and imminent—say, tomorrow a gamma-ray burst from our solar system will strike Earth and kill all life—then the plausibility leg of the tripod is too short and so the tripod does not stand, there is no supreme emergency, and no fast science is warranted (low probability threats of catastrophe should of course be studied by routine science). Yet, if the three legs of the threat tripod are tall enough, the supreme emergency justification says: catastrophe is very likely coming soon and so we should engage in and act on fast science.

The supreme emergency justification asks us to forgive science for being more hare than tortoise, and to put aside misgivings about the potential harms, both epistemic and practical, of violating the norms of routine science. Yet, the supreme emergency justification is incomplete, because it says nothing about how hare-like fast science can be. Here lies a worry, because, as above, fast science departs from routine science by violating the practices that render routine science reliable. Scientists are invited to give up on some of those good-making features of science. But this cannot be an invitation for scientists to consult tea leaves, tarot cards, and mystics. There must be limits to fast science, especially in a context in which policymakers claim to 'follow the science'. My aim in this paper is to articulate and defend two principles for guiding and assessing fast science.

The first I call the Principle of Similarity, which says that fast science should be as similar as possible to routine science while satisfying the temporal and material constraints of the supreme emergency that motivate the fastness of the fast science in the first place. The supreme emergency warrant for fast science does not entail that anything goes, but rather that scientists can cut some corners of routine science while staying as close as possible to the reliability-enhancing practices of routine science. This principle is straightforward, and as I note in §3, is so basic that it needs little argument (though in §4 I describe a recent case in which the principle was to some degree flouted). The second I call the

Principle of Independence, which says that violations of the reliability-enhancing practices of routine science when developing mitigation strategies for a threat should not be the same or similar to violations of the reliability-enhancing practices of routine science that were used to detect the threat (in §3 I articulate this principle in more detail and explain what 'same' or 'similar' means in this context). That is, scientific work which is intended to inform our choice of intervention should not depend on the same problematic assumptions or methods which were used to estimate the magnitude, imminence, and plausibility of the threat itself.

Similarity and independence are in general graded notions. An instance of fast science satisfies the two principles to some degree or other. And of course the magnitude, imminence, and plausibility of a threat are graded properties—the taller the threat tripod, the greater the supreme emergency context of fast science. So, a supreme emergency involves considering graded properties that speak in favour of the violation of routine science practices (the threat tripod) and graded properties that speak against that violation (the two principles). This evaluation can occur both *ex ante* and *ex post*. Assessment *ex ante* could take the scientist's perspective and ask: given the estimation of the threat tripod, to what degree ought I violate the principles and practices of routine science? Assessment *ex post* could take the decision-maker's perspective and ask: given the estimation of the threat tripod, to what extent should my choice of intervention be based on results of scientific work which violated principles and practices of routine science? (Assessment *ex ante* could also take the policy-makers perspective and assessment *ex post* could take the scientist's perspective.)

After further characterising fast science in §2, I defend the two principles for fast science in §3. I describe two famous episodes of fast science in §4—the Manhattan Project and the Covid-19 epidemiological modelling—and I use the two principles to assess these episodes. Contexts of fast science are also contexts of decision under uncertainty, and so in §5 I describe how fast science relates to theories of decision under uncertainty and great threat, particularly versions of a precautionary principle. In §6 I conclude that fast science cannot be a tortoise, though ought not be hare.

2. Fast Science

Although there is an obvious temporal dimension to fast science, it is not merely the pace of scientific work that distinguishes fast science from routine science. Much routine scientific work is carried out in conditions of great temporal pressure resulting from priority races and the tenacity of ambitious scientists, as occurred with the discovery of the double helical structure of DNA by Watson and Crick. The reward structure of routine science can motivate hasty work (Heesen [2018]). Conversely, some instances of scientific work that have many of the features of fast science occur over a period of several years, as occurred during the Manhattan Project (an example I return to below). Rather, what distinguishes fast science from routine science is the violation of reliability-enhancing principles and practices of routine science (whatever those are), in response to a supreme emergency.

There is a cluster of principles and practices that render routine science relatively reliable. Here I am agnostic as to whether Merton's norms ([1942]), Longino's criteria ([1990]), Kitcher's conditions ([1982]), or some other set of norms are the markers of genuinely good science. Moreover, particular reliability-enhancing practices are domain-specific: testing the effectiveness of a new drug might require randomised trials, while predicting the future location of an astronomical object might require simulations; some scientific communities are organised around practices such as peer review, while others share findings via online repositories. These domain-specific practices themselves vary on even more fine-grained properties—for example, a randomised trial can involve one hundred subjects or one hundred thousand subjects.

To understand fast science we do not need a complete list of all the reliability-enhancing practices of routine science, and nor do we need a single grand principle of demarcation, as we do not need to suppose that the distinction between fast science and routine science is sharp—it is enough to say that routine science has practices which minimise epistemic risks and enhance the reliability and objectivity of science (Koskinen [2020]), and fast science is action-guiding science which departs from some or many of those reliability-enhancing practices and receives putative warrant for that departure based on the threat tripod of a supreme emergency.

There is a recent movement dubbed 'slow science'. Slow science is characterised by careful, methodical work, motivated by curiosity rather than practicalities, and resists the publish-or-perish model of routine science (Stengers [2018]). Though fast science and slow science naturally differ in many important respects, fast science is best characterised by reference to routine science rather than slow science.

The fastness of fast science—and to repeat, by this I do not strictly mean the temporal dimension—can occur at two distinct phases. The first I will call the 'detection' phase, which is the scientific work that provides estimates of the magnitude, imminence, and plausibility of a threat. An assessment of whether a scenario is a supreme emergency may depend on fast science in the detection phase. The second phase I will call the 'intervention' phase, which is the scientific and technical development of threat mitigation strategies. Considering the cases I discuss in §4, in the Manhattan Project these two phases were temporally and scientifically distinct, while in the Covid-19 modelling during the spring of 2020 these phases were to some extent collapsed into the same work and resulting Report 9 (Ferguson et al. [2020]).

Some scientific work is similar to fast science in that it violates some of the principles and practices of routine science, but without the supreme emergency context. Elliott and McKaughan describe the scientific assessments of restored wetlands ([2014]), which must be 'authoritative, cheap, and quick,' are not peer-reviewed or formally published, and are policy-guiding (see Steel [2016] for a criticism). Expedited chemical screening is another domain in which the trade-off between speed and accuracy is influenced by non-epistemic considerations. While the practical context of assessing restored wetlands and expedited chemical screening may provide a putative justification for not bothering with the routine science principles and practices, the practical context is clearly not *in extremis*. Appeals to science in legal contexts are often similar (see for example Miller [2016]). In this paper I am concerned only with fast science in response to catastrophic, imminent, plausible threats.

Fast science can be illuminated by considerations of inductive risk. The inductive risk argument holds that there are two basic ways a scientist can err: she can accept a hypothesis when it is false and reject a hypothesis when it is true. Both kinds of errors have practical consequences, and, for a particular hypothesis, our evaluation of those consequences influences our evidential thresholds for accepting or rejecting the hypothesis and acting accordingly (see, for example, Rudner [1953], Douglas [2000], Steele [2012], and Biddle and Kukla [2017]). In some scenarios we might be willing to accept and act upon a hypothesis with very little support because the consequences of not acting could be catastrophic. There are three features of fast science that render inductive risk considerations particularly acute. First, the uncertainties of fast science are typically greater than the uncertainties of routine science—yes, routine science has inductive risk, but that is in the context of scientific work that has the reliability-enhancing feature of routine science, while fast science by definition involves the violation of some of those reliability-enhancing features and hence an increase in uncertainty. Second, one of the usual responses to the inductive risk argument in routine science is to conclude that values influence scientific reasoning, and thus scientists must carefully manage that value-influence, by, for example, being transparent about the influence of values (Elliott [2022]), being selective about which values are involved (Kourany [2010]), appealing to democratic values (Schroeder [2021]), developing shared conventions (Wilholt [2013]), and enriching the diversity of value-perspectives (Longino [2004]). That is all useful. But it all takes time. Routine science has that time. The fastness of fast science might hinder these value-management strategies. Third, the threats in fast science are catastrophic in magnitude, while the risks in routine science can include harms of any magnitude, which may be catastrophic (say, if our climate models severely underestimate the impact of carbon emissions), yet many such risks are, in comparison to catastrophe, minor in magnitude.

In a recent insightful article, Friedman and Šešelja argue that higher order evidence can be important in assessing scientific controversies, particularly in the context of fast science ([2023]). By fast science Friedman and Šešelja mean 'application-driven research confronted with an urgent need to accept or reject a certain hypothesis for the purposes of policy guidance, aimed at addressing a significant pending social harm' ([2023], p. 938). Important for their definition, delaying assessment of a hypothesis of concern to allow time to gather more evidence may be harmful, and hence the context of fast science requires endorsing (or rejecting) a hypothesis without waiting for further evidence. That is consistent with my account of fast science, as a typical possibility in routine science is to gather more and better evidence for a hypothesis, and not doing so may amount to violating a reliability-enhancing principle of routine science. The question I address here is about the putative justification of fast science, while the question Friedman and Šešelja address is about the importance of higher-order evidence for thinking about how to respond to scientific disputes during fast science.

Shaw defines 'urgent science 'as follows: 'A research proposal is urgent iff there is a practical or moral reason to need a result within a specified timeline and the research can realistically be carried out within that timeframe.' ([2022a], p. 108; see also Shaw [2022b] for a discussion of criteria which can be used to assess whether a research project is urgent). Shaw articulates this important idea in the context of discussing pursuit-worthiness of scientific research programs. The notion of urgent science is clearly related to my notion of fast science, as fast science gets its putative warrant based on a practical or moral reason to need quick results. Yet, Shaw's notion of urgent science is much broader than my notion of fast science, as I am specifically concerned with scenarios of supreme emergencies.

Like a spy in the dark alleys of an old European town, threats can sneak up on you. Information about the magnitude, imminence, and plausibility of a threat can trickle in via various channels that look a little like routine science, until you wake one morning to find yourself in a supreme emergency. Developments in atomic physics in the late 1930s culminated in physicists recognising that atomic energy could be harnessed to build massive bombs and that the Germans were probably pursuing this. Similarly, the first three months of 2020 involved a trickle of reports about the spread and lethality of the new coronavirus which culminated in the dire predictions of the epidemiological modelling in March 2020 (White et al. [2022]). The gradual sneaking up of a threat explains why the detection phase of fast science might require violating some routine scientific practices. The fastness of the intervention phase is, in turn, putatively warranted by the supreme emergency justification and by the fact that, based on the novelty of the threat, there is little existing routine science results which can be relied on for developing interventions.

One might ask if there is a potential circularity regarding the justification of the detection phase. The task of the detection phase is to assess the threat tripod, but the establishment of the threat tripod is precisely what justifies the violation of reliability-enhancing principles and practices of routine science. Yet violations of the reliability-enhancing principles and practices of routine science in the intervention phase plausibly requires greater justification with respect to the threat tripod relative to the detection phase, since the intervention phase has implications for policy while the detection phase does not itself have direct implications for policy. Both phases of fast science are literally phases and not moments—above I noted that supreme emergencies can sneak up on us; both the lead-up to the Manhattan Project case and the early months of the Covid pandemic exemplified this. The earliest signals of a threat can be (and perhaps often are) nebulous, and can suggest a threat in a manner without all three legs of the threat tripod standing tall, but this can anyway warrant further work in the detection phase. The supreme emergency justification for acting on the results of fast science from the intervention phase, however, requires the full threat tripod.

To summarise, fast science:

i. Violates some or many of the reliability-enhancing practices and principles of routine science at its best
ii. Is meant to guide policy responses to a supreme emergency
iii. Can be characterised as having two phases: the detection phase and the intervention phase

The supreme emergency justification adds: (i) is putatively justified by (ii). One might note that some instances of routine science, unfortunately, also have feature (i). If so, we may not be able to easily demarcate instances of fast science from instances of bad science. If one were motivated by this aim of demarcation, one could simply assess whether feature (ii) is present in addition to feature (i). In what follows I ask about the putative supreme emergency justification of fast science.

3. Similarity and Independence

Consider this case:

> Asteroid Impact
> Maria consults astronomical observations and established theory, and predicts that an asteroid will soon strike the earth, causing incredible destruction. In a panic, she does some quick back-of-the-envelope calculations to ascertain the most effective mitigation strategy. She writes a report for policy-makers and holds a press conference urging immediate implementation of her recommended intervention.

In the intervention phase of Asteroid Impact Maria's method departs from routine science. Maria violates the Principle of Similarity.

The Principle of Similarity is intuitive, and indeed is so straightforward that it hardly needs an argument. It is worth stating, however, since, as we will see in §4, some prominent examples of fast science did not adequately satisfy the principle. If a threat's magnitude, imminence, and plausibility are sufficiently severe, we might be willing to forego the reliability-enhancing practices of routine science, but there are limits: the scientific work in fast science, both in the detection phase and particularly in the intervention phase, ought not depart so radically from routine science such that the resulting work becomes highly unreliable. If such work does depart so radically from routine science, it ought not be relied on for policy. Asteroid Impact involves methods that depart from routine science. Yet similarity is a graded notion. Any reason for routine science to deploy a practice on the grounds that it is reliability-enhancing is also a defeasible reason for an instance of fast science to deploy that practice. The very idea of warranted fast science supposes that supreme emergencies provide at least some justification for the violation of at least some of the practices of routine science, yet the standing reasons for those practices remain and ought to constrain the resulting fast science methods as much as possible, within the temporal and practical bounds of the threat tripod.

Consider a broad principle, say, Longino's norm that science should be responsive to criticism ([1990]). An episode of fast science could involve a group of scientists with no critical feedback internal to the group, and present their work to policy-makers without first seeking criticism from external experts, and subsequently be impervious to criticism—such fast science would be maximally dissimilar to routine science on this broad norm of responsiveness to criticism. Another group doing similar work could include some practice of critical feedback internal to the group and seek criticism from external experts, and such fast science would be more similar to routine science on this norm, and thus would better satisfy the Principle of Similarity.

The Principle of Similarity can also be assessed by reference to fine-grained reliability-enhancing practices. Suppose that to properly test a new drug in a clinical trial in a routine science context 100 subjects would be needed. Suppose that a threat tripod entails that we do not have time to test 100 subjects: a trial with 99 subjects would satisfy the Principle of Similarity more than would a trial with 98 subjects, and a trial with 98 subjects would satisfy the Principle of Similarity more than would a trial with 97 subjects, and so on.

A referee notes that it may be unclear how the Principle of Similarity can guide action in concrete cases. Many philosophical principles, though, are not always clearly action-guiding, and yet the project of articulating those principles can be nonetheless worthwhile. The principle of expected utility maximisation, for example, has this property (since determining the requisite probabilities and utilities to assess the expected utilities of outcomes may be impossible in many practical contexts), but the principle is anyway worth having and can guide action in at least some cases.

Walzer famously argued that supreme emergencies can justify the 'dirty hands' of political and military leaders who act immorally in order to save their community, by, for example, bombing civilians or torturing informants ([1973]). Birch applied Walzer's argument to conclude that scientists can justifiably violate scientific norms in supreme emergencies ([2021]). Walzer's argument is controversial; torture, for example, is unreliable, and there is a potential incoherence to the thought that immoral behaviour can be morally justified. In any case, there is an important difference between the supreme emergency justification of engaging in and acting on fast science and the supreme emergency justification of dirty hands. The justification of dirty hands is supposed to warrant behaviour that is drastically *unlike* the sort of behaviour that is permitted by routine moral principles, and more drastic departures from moral behaviour may be more effective in achieving one's aim. Yet in supreme emergency scenarios, the more that fast science departs from the principles and practices of routine science, the less reliable is the science, and so the recommended intervention would in turn be less likely to be effective. When the legs of the threat tripod are sufficiently tall, immoral behaviour may be warranted (though that is controversial), but unscientific behaviour is not warranted.

Let us now consider the second principle for assessing fast science. The Principle of Independence says: scientific work which guides our choice of intervention should not involve the same violations of the reliability-enhancing practices of routine science which were used to assess the threat tripod, so long as it does not come at the cost of decreased reliability. I offer two distinct arguments for the Principle of Independence. The first is based on considerations of reliability, like the argument for the Principle of Similarity. The second is based on a second-order principle that holds that violations of norms cannot be self-justifying.

The Principle of Independence is meant to guarantee that whatever specific form of unreliability exists in the detection phase of fast science is not inherited by the intervention phase. Recall that what defines fast science is the violation of some reliability-enhancing principles and practices of routine science. We cannot appeal to the potential unreliability of a method in an instance of fast science as grounds to not attend to its results or act on those results, since the reliability of all fast science ought to be questioned—because it is fast science, some of the reliability-enhancing principles and practices of routine science are violated—and yet we have granted the supreme emergency justification of engaging in and acting on at least some fast science. It follows that there may be an estimation of a severe threat tripod in the detection phase, yet even if the Principle of Similarity is satisfied, the fast science in the detection phase could be unreliable in a particular way, and that particular form of unreliability could be inherited by the intervention phase. The resulting errors in estimating the threat tripod could be compounded by the same errors in assessing mitigation strategies. The Principle of Independence is meant to guarantee that this does not occur.

The Principle of Independence holds that, so long as it does not come at the cost of decreased reliability, scientific work which guides our choice of intervention should not depend on the same violations of the reliability-enhancing practices of routine science which were used to assess the threat tripod. Yet, 'same' here is ambiguous: it could mean the same token violation or the same type of violation. In §4.2 I give an example in which the same token violations of reliability-enhancing practices occur in both the detection phase and the intervention phase. In the coming arguments I focus on token-sameness, as it is particularly problematic, yet the arguments that follow also apply, though perhaps less forcefully, to type-sameness.

To illustrate the importance of the Principle of Independence, consider the next hypothetical case:

> Pandemic
> Lilia simulates the spread of a new infectious disease using her in-house software, which has been developed over years and informed by past epidemics, and has been the basis of many peer-reviewed publications. She uses the scant data she can access to set the values of the parameters in the simulation. Her results suggest that the disease will cause terrible catastrophe. Using the same model, she also simulates mitigation strategies, and her results suggest that her favoured strategy will save millions of lives. She writes a report for policy-makers informing them of her findings and holds a press

> conference urging immediate implementation of her recommended intervention.

The fast science in Pandemic satisfies, as much as could be possible given the supreme emergency context, the Principle of Similarity, in both the detection phase and the intervention phase. In that regard, this case is superior to Asteroid Impact. Yet, unlike Asteroid Impact, Pandemic involves the very same token methods, assumptions, and potential sources of unreliability (particularly the scant data used to set parameter values) in both the detection phase and the intervention phase, and so the Principle of Independence is not satisfied. Now suppose that the scant data that Lilia uses to set her parameters are unreliable, and in particular, those parameter values cause the simulation to grossly overestimate the speed of disease spread and the lethality of the disease. That overestimation, in turn, contributes to a gross overestimation of how many lives would be saved if her favoured mitigation strategy were implemented. This is not a departure from the Principle of Similarity, since the work is as close as possible to what could have been achieved in routine science given the temporal constraint. The problem is that the specific source of unreliability in the intervention phase—namely, the value of the simulation parameters—is the same token source of unreliability in the detection phase. Were the Principle of Independence satisfied, that would have been avoided.

This rationale is similar to diversity of evidence arguments discussed in philosophy of science, which holds roughly that the more diverse is a set of evidence supporting some hypothesis, the more likely that hypothesis is true (see Kuoriksoki and Marchionni [2016], Schupbach [2018], and Claveau and Grenier [2019]). Yet in those arguments the multiple lines of evidence are about the same hypothesis or phenomenon, while the two phases of fast science are about two distinct phenomena: the estimation of the threat tripod in the detection phase and properties of mitigation strategies in the intervention phase. There is a risk of error in both the detection phase and the intervention phase, and just as it is good to spread risk across multiple investments in a portfolio, it is good to spread risk across the two phases of fast science. (The investment portfolio analogy helps to show that token-sameness error risk is worse than type-sameness: a very risky investment strategy is to invest all of one's resources into, say, shares of one particular technology company, while a strategy which is less risky, though risky nonetheless, would be to invest in multiple token technology companies; yet that latter strategy involves investing in a single type of company, so a still less risky strategy would be to invest in multiple types of companies from different industries.)

When the Principle of Independence is satisfied, the results of the intervention phase can often serve as a robustness check on the results of the detection phase. Since we have granted that some reliability-enhancing features of routine science can be violated in both phases, such a robustness check can be valuable in minimising error. Suppose, for example, that in the case *Pandemic*, in the intervention phase researchers did not rely solely on the model to evaluate interventions but rather in addition they performed a cluster randomised study in which some communities were randomly allocated to particular mitigation interventions and others randomly allocated to no intervention; the results of the intervention phase, which satisfies Independence, would serve as an independent test of the results of the detection phase. That is because the control group of such a study would provide evidence relevant to the predictions that the model made about a scenario in which policy-makers do nothing, and the intervention groups of the study would similarly provide evidence relevant to the various predictions made by the model in the various intervention scenarios. If the model in the detection phase predicted that roughly x number of people will get infected with no intervention but roughly y number of people will get infected if we close schools, and if an extrapolation from the trial roughly confirms those values of x and y in the two respective scenarios, then policy-makers would have much better epistemic warrant for their decision to either close schools or keep them open. Yet if the Principle of Independence is not satisfied, such a robustness check is not possible, since both the detection phase and the intervention phase share the same risks of error. (One might worry that performing a randomised trial would take too much time in the context of a supreme emergency; the point here is simply to articulate the epistemic benefits of satisfying the Principle of Independence.)

Here is a slightly different way to make this point. The fastness of fast science is putatively justified by a supreme emergency. Estimation of the threat tripod occurs in the detection phase. The results of the detection phase, in turn, offer putative justification to the fastness of the intervention phase. Yet, the detection phase has risk of error, and so the justification of the fastness of the intervention phase is threatened. If the intervention phase does not share the same risks of error as the detection phase, then the results of the intervention phase can provide independent justification for the results of the detection phase; since it is the results of the detection phase which provides justification for the intervention phase, the result is a mutual justification between phases, when Independence is satisfied. If Independence is violated, on the other hand, this mutual justification is not available.

Let us turn to the second argument for the Principle of Independence. Each of the many general principles and fine-grained practices of routine science that enhance its reliability can be formulated as a corresponding *pro tanto* norm. Here are some examples: insofar as randomisation renders trials more reliable (Larroulet Philippi [2022]), the corresponding norm is, randomise; insofar as absolute measures of effectiveness are more reliable than relative measures for choosing treatments (Jantgen [forthcoming]), the corresponding norm is, report absolute measures; insofar as modulating inferences to account for biases renders those inferences more reliable (Erasmus [forthcoming]), the corresponding norm is: modulate inferences to account for biases. Fast science violates norms of routine science. Since they are *pro tanto* norms their violation may be justified, perhaps on grounds of feasibility, cost, ethical constraints, or a supreme emergency. Yet the violation of any norm, *pro tanto* or not, scientific or not, fast science or not, cannot be justified by the very instance of norm violation needing justification. Norm-violation cannot be self-justifying.

The viciousness of self-justification of norm violation may be more vivid in non-scientific contexts. Imagine Lilia is breaking the speed limit as the only driver on a road passing a school, and a police officer pulls her over and tells her she was breaking the law by speeding, and she responds by saying that she was just driving at the average speed of all cars on the road at the time. Imagine Neil is caught meeting his lover during a lockdown in which non-household social contacts are prohibited on grounds that they could cause harm, and he responds by saying that his meeting was fine as it brought him a lot of pleasure. Imagine a devoted Kantian who sincerely believes that lying is never permissible and then lies about the permissibility of lying by appealing to its utility-maximising consequences only to justify to himself that very lie. To be clear, Lilia's speeding, Neil's tryst, and even the Kantian's lying may have been justified violations of norms, yet those justifications must be based on something other than the very act in question. Perhaps Lilia could cite the fact that she was bringing her ill child to a hospital. That would not be a self-justification of norm-violation, as it would appeal to something other than the act itself or features of the act itself as grounds for violating the norm (see Gert [1998], Chapter 9, for a discussion of norm violation).

Here is a general statement of what I am concerned with:

> Self-justification of norm-violation
> *S* does action *A* which violates norm *N*, and *S* cites features of *A* as justification for violating *N*

Features of *A* that one might cite (absurdly) to self-justify a norm-violation could include its statistical normality (Lilia's speeding), the goods that it brought to *S* (Neil's tryst), or its sheer permissibility (the Kantian's lying). In each case the putative self-justification is risible. To be sure, citing features of *A* can be relevant to assessing *N*'s status as a norm. Suppose *A* resulted in no harm, and a compelling case could be made that actions like *A* rarely result in harm but they bring many benefits. This could challenge the status of *N* as a norm, though it would not justify *S*'s behaviour in that instance; moreover this consideration is not applicable to fast science, since fast science involves violating a reliability-enhancing practice and not challenging whether that practice is indeed reliability-enhancing. Self-justification of norm violation can be warranted when the positive consequences for doing so are great, as consequentialists have long argued (Gert [1998]); in the context of fast science this places an important constraint on the Principle of Independence, namely, that if violating the principle increases the reliability of the intervention phase, one can or ought to do so.

Let us consider a scientific example of self-justification of norm-violation. We can imagine a scientist who does *A* which violates *N*, and when challenged about her violation of *N*, she cites features of *A*, particularly the outcome of doing *A*. For example: she is testing a new fertiliser for sunflower plants, and her statistician tells her that to get sufficient statistical power she will need to randomly allocate 1000 sunflower plants to a fertiliser group and a control group; yet after getting data on only 20 sunflower plants, she stops the trial, analyses the data, and gets results *R*, which suggests that the sunflower plants grew taller than she expected with the fertiliser. The statistician criticises her, stating that because her trial did not include a sufficient number of sunflower plants, *R* is unreliable; the scientist responds by citing the large effect size of *R* as grounds for violating *N*. Yet the statistician considers *R* to be unreliable precisely because it violated *N*, and thus considers *R* no grounds at all for violating *N*.

The self-justification of norm violation in the cases above are synchronic: *S* cites features of *A* as justification for the violation of *N* by that very act *A*. Yet, fast science is defined as occurring in two stages, and if those stages are temporally distinct then the self-justification of norm violation would have to be diachronic: *S* cites features of A_1 at time T_1 as justification for the violation of *N* via A_2 at T_2. In response, note that the cases above can be given a diachronic structure. Neil could have met his lover on several occasions, and on the occasion in which he his finally caught he could offer as putative justification the pleasure that his prior trysts brought him. That would be as absurd as the synchronic cases above. (Also, as we will see in §4.2, in the case I use to illustrate the violation of Independence the self-justification of norm violation is in fact synchronic.)

It is worth noting that identifying which norms have been violated may not always be obvious, and in turn it may not always be easy to discern if the Principle of Independence is satisfied. Some violations of reliability-enhancing principles and practice of routine science are transparently discernible while others are harder to discern.

I have conceived of fast science as composed of two phases: estimation of the threat tripod and assessment of optimal interventions for the threat. The Principle of Independence states that violations of the reliability-enhancing practices of routine science in the intervention phase should not depend on the same violations in the detection phase, so long as this does not come at the cost of decreased reliability. In §1 I argued for the supreme emergency justification for violating the reliability-enhancing practices of routine science, including in the detection phase. However, violations of reliability-enhancing practices in the intervention phase cannot be given the supreme emergency justification if the assessment of the threat tripod was based on the same violation of reliability-enhancing practices in the detection phase, as that would amount to self-justification of norm-violation.

To be clear, both principles are important only as means to minimise error and maximise the chance of good outcomes from whatever policy is deployed. They are defeasible, particularly given a terrible threat tripod—recall in §1 I argued that the magnitude, imminence, and plausibility of a threat are all graded properties of a supreme emergency, and the greater they are, the more the two principles defended here can be violated. One might interpret the aim of the two principles as something like 'minimise error, subject to the time constraints imposed by a supreme emergency.' Yet I am open to the possibility that there may be other relevant principles, particularly of a non-epistemic nature, depending on the details of a supreme emergency. For instance, suppose a supreme emergency involves a threat to human reproductive capacities, and the birth rate plummets to near zero. Scientists might then do all they can to maximise the probability of discovering an effective intervention and given the nature of the threat they might decide to care little about false positive discoveries and care immensely about false negatives, on broadly inductive risk considerations—and so maximising the chance of a good outcome might involve increasing the frequency of a particular kind of error (false positives).

There may be an interaction between Similarity and Independence: if Similarity is satisfied to a very large degree in the detection phase, then it might not be so bad to violate Independence in the intervention phase (because the detection phase methods were pretty reliable), and indeed it might be more reliable to violate Independence compared with using a very unreliable method which satisfies Independence; but on the other hand, if Similarity is grossly violated in the detection phase, then it is very important to satisfy Independence in the intervention phase (precisely so that the unreliability in the detection phase does not carry over to the intervention phase).

4. The Tortoise and the Hare

In this section I describe in more detail two episodes of fast science referred to earlier. The first is more tortoise than hare, the second is more hare than tortoise.

4.1 Manhattan Project

For an example of fast science that arguably satisfies both principles we can turn to the Manhattan Project. Although it has been described as 'big science' (Hughes [2003]), the Manhattan Project also had many features of fast science: the work was done with extreme secrecy and thus little by way of usual peer review of scholarly publications, performed under extreme time pressure, with a clear action-oriented goal. The magnitude and imminence of the threat—Germans being first to develop atomic bombs—became clear to emigre physicists by the late 1930s, leading to the Einstein-Szilard letter warning President Roosevelt about the the possibility of 'extremely powerful bombs of a new type.' This was the detection phase.

During the Manhattan Project many scientists were opposed to the secrecy and compartmentalisation of their work. For example, in 1944 Szilard wrote a letter to Vannevar Bush in which he stated that 'decisions are often clearly recognised as mistakes at the time when they are made by those who are competent to judge, but … there is no mechanism by which their collective views would find expression or become a matter of record' (cited in Rhodes [1986], p. 508). When Teller tried to convince Bohr that nuclear research should be kept secret during the war, Bohr responded by insisting that secrecy should not be a part of science (Rhodes [1986], p. 294). Contrary to routine science, atomic physics of the Manhattan Project involved no peer review of the usual sort, no publication, extreme secrecy, all in the service a military technology.

The intervention phase was the Manhattan Project itself. The Manhattan Project included an intelligence component which was tasked with gathering information about German atomic science and technology, but that work was substantively independent from the scientific and technical developments of the Manhattan Project. Obviously results from the espionage could have contributed to the scientific and technical developments of the Manhattan Project, and the scientific and technical developments of the Manhattan Project could have influenced the estimation of the magnitude, imminence, and plausibility of the threat. For example, failure to enrich uranium despite sustained effort would have provided some evidence that the Germans were also having difficulty enriching uranium, and that in turn would influence the threat tripod, while success at enriching uranium would have had the opposite consequence—that additional evidence about the threat tripod would be especially compelling given how different the nature of the work was in the intervention phase compared with the detection phase. Whatever violations of the principles and practices of routine science occurred during the intervention phase, those were different from whatever violations of the principles and practices of routine science occurred during the detection phase. The Principle of Independence was satisfied.

The scientific work was informed by decades of prior nuclear science, and was carried out by many of the most eminent physicists of the day. The director of the Manhattan Project, Robert Oppenheimer, convened a conference devoted to discussing nuclear weapons in 1942. Huge sums of money were devoted to empirical, theoretical, and technical work. This work was distributed over many research sites, and was subject to immense critical scrutiny, particularly between different teams. In short, though there were departures from the practices of routine science, the Manhattan Project involved reliability-inducing practices that served similar functions as those routine practices, and so the Principle of Similarity was to a large degree satisfied.

4.2 Covid Models

The Imperial College London Covid-19 modelling and resulting Report 9 made dire predictions about the number of deaths and the number of intensive care beds as a result of the spread of the Covid-19 coronavirus, under a variety of scenarios which ranged from doing nothing to using a range of mitigation strategies (Ferguson et al. [2020]). Birch ([2021]) notes that Report 9 urged the deployment of a particular policy, namely sustained lockdown, as 'the only viable option' (10) (Report 9 uses the phrase 'the preferred policy option'). As noted in §1, this work lacked many of the reliability-enhancing practices of routine science. Critics of that fast science and the resulting policies argued that crucial parameters of the model were poorly substantiated, and that further sensitivity analyses would have demonstrated that the dire predictions were not robust to changes to those parameters (see, for example, Winsberg et al. [2020), a response by van Basshuysen and White [2021], and a subsequent rejoinder by Winsberg et al. [2021]). Two of the most important parameters were the reproduction number and the infection fatality rate, and while there was already empirical evidence available at the time which suggested that the Imperial College London values for these parameters were probably inaccurate, subsequent work confirmed this (Northcott [2022]). As noted earlier, Report 9 was not peer-reviewed, and was presented directly to policy-makers. Not long after its publication the projections of Report 9 appeared to be grossly inaccurate (Winsberg et al. [2021]). In short, the Principle of Similarity was, to a very significant degree, violated.

Why didn't the scientists include a richer set of data to inform their parameters, and why didn't they perform more sensitivity analyses? A plausible answer is that they were under time pressure given the nature of the threat, and they simply could not wait for better data to inform their parameters or to perform more simulations. A delay in reporting their findings to policy-makers could have cost many lives. Yet specific predictions about the threat tripod and how to respond to the threat were the entire point of Report 9. Consider, for example, the data in Table 4 of the report, which suggests that in the United Kingdom the number of deaths over two years would be about half a million greater if no mitigation policies were deployed. Winsberg ([2022]) suggests that this modelling work was 'self-recommending', insofar as were one to insist that the scientists should gather more evidence to inform the model, this would take time, and thus many people would die, as the model predicts … and thus they need not gather more evidence. That self-recommending aspect of this work constitutes a violation of the Principle of Independence.

At the time that the Imperial College London epidemiologists presented Report 9 to policy-makers—mid-March 2020—there were compelling reasons to infer that the threat of the pandemic was large, imminent, and plausible, particularly based on experiences in China and Italy. Yet, the conclusions of Report 9 were profoundly dire, and presented with extreme precision and certainty, in an authoritative context—the lead scientist, Neil Ferguson, had been a member of the Scientific Advisory Group for Emergencies (SAGE) from its first meeting about the pandemic in January 2020. Report 9 itself contributed specific and authoritative estimates of the magnitude, imminence, and plausibility of the threat of the Covid-19 pandemic. Using the very same model, with all its shortcomings, Report 9 provided very specific guidance on mitigation strategies. Figure 2 in Report 9 is a visual depiction of the simultaneous assessment of the threat tripod and prediction of the effectiveness of various mitigation strategies, insofar as it includes simulation results for both 'do nothing' scenarios and those with mitigation (page 8 of doi.org/10.25561/77482). This is a violation of the Principle of Independence.

This might be a common feature of epidemiological models. van Basshuysen et al. ([2021, p. 119]) note:

> Epidemiological models serve dual purposes: apart from their epistemic purpose of forecasting the course of an epidemic, they also serve the practical purpose of informing and guiding policy-making. These epistemic and practical purposes go hand in hand: on the basis of forecasts, policy-makers can choose policies that are likely to prevent unwanted outcomes.

That seems right to me. Yet if an epidemiological model has a significant flaw, then both of its dual purposes could be hindered by that flaw, and the resulting policy could be significantly misguided, and the concern about self-justification of norm-violation stands. Fast science, including epidemiological modelling, should strive to satisfy the Principle of Independence.

***

These two cases are merely meant as illustrations, and it would be helpful for future work to explore a richer range of cases. The development of the Russian Sputnik V Covid-19 vaccine, for example, is plausibly a case in which Similarity was violated, since the vaccine was rolled out in Russia after only a phase 2 trial rather than the usual phase 3 trial (and Sheldrick et al. [2022] re-analysed the data from the trial and concluded that data showed signs of manipulation), but this work plausibly satisfied Independence (since the violation of the reliability-enhancing practices in the detection phase simply had nothing to do with trials). It would be interesting for future research to explore a wider range of cases of fast science.

5. Caution and Intervention

Suppose there is a supreme emergency and the resulting fast science satisfies both the Principle of Independence and the Principle of Similarity. Does that entail that whatever mitigation strategy is recommended in the intervention phase is best? It does not. Conversely, suppose the Principle of Independence and the Principle of Similarity are not satisfied in the intervention phase. Does that entail that whatever mitigation strategy is recommended in the intervention phase is not best? It does not. Fast science is still just science—it is not an inference rule or decision procedure. Yet clearly the results of fast science could be relevant to decisions in contexts of uncertainty and great threat.

Like any decision under uncertainty, there are competing views about how agents should choose among various options. Some hold that the choice to intervene in a particular way should be assessed by a full cost-benefit analysis, even if estimates of the probabilities of outcomes are unavailable—Broome, for example, advises us to 'stick with expected value theory' ([2012], p. 129), and Posner advises standard cost-benefit analysis even in the face of a catastrophic threat ([2004]). Some hold that the best approach is to postpone a decision until more information is available (see, for example, the discussion in Bradley and Steele [2015], §4.2). Some advise deciding cautiously, by aiming to avoid the worst possible outcomes. In this section I ask how fast science relates to theories of decision under uncertainty.

Fast science can obviously give us information about a threat and inform policymakers about options for intervention and their chances of success. Yet, Birch claims that fast science can do more ([2021]). He claims that *in extremis* contexts can warrant an 'imperative to elicit a specific policy response' (p. 90), because if that response were not elicited the advisor believes catastrophe would result. However, results of fast science grossly underdetermine the best intervention. The mitigation strategy recommended in the intervention phase of fast science could be extremely costly, or unfeasible, or have profoundly terrible unintended but foreseeable consequences, or have a low probability of success; the fastness of fast science entails that we should have some uncertainty about the probability of intervention success; as Birch rightly notes, science advisors are not accountable to an electorate as democratically-elected politicians are; finally, science advisors have interests and values that are likely to be unrepresentative of the broader public. To use Pielke's terms, scientists should be 'honest brokers' rather than 'issue advocates' ([2007]) (though there has been some debate about Pielke's honest broker ideal; see Havstad and Brown [2017]; Brown and Havstad [2016]).

Perhaps a more robust connection between fast science and policy response can be grounded on a precautionary principle, given that precautionary approaches are typically recommended when, and only when, there is a plausible catastrophic threat.

At the most general level, precautionary principles appeal to estimates of the magnitude of a threat and estimates of the costs of intervening on that threat and recommend avoiding worst-case scenarios. Yet because there are various versions of the precautionary principle defended in the literature, there is little to say about the general relation between fast science and a precautionary approach to decision under uncertain and extreme threat.

Harnessing Rawls's maximin principle, Gardiner ([2006]) argues that precautionary interventions are warranted in response to a potential catastrophe when these four conditions are met: (i) there is no information about the probabilities of the possible outcomes or what little information there is ought to be sharply discounted, (ii) the intervention guarantees that the catastrophic outcome would not occur, (iii) the cost of the intervention is minimal, and (iv) alternatives to the intervention, including doing nothing, may result in the catastrophic outcome (see also Steel [2015] for discussion). In a supreme emergency, arguably (i) would be met (on grounds that the emergency constrains the amount of time science could devote to getting sufficient evidence to inform the probabilities). I reckon that (iii) would rarely be met in supreme emergencies, but in any case (iii) is more applicable to regulatory contexts and can plausibly be discarded in a supreme emergency (Sunstein argues that it should be discarded in any context ([2005], p. 382)). Condition (iv) could be met. However, I reckon (ii) could rarely be met in supreme emergencies, and even if both the Principle of Similarity and the Principle of Independence were maximally satisfied, the residual uncertainty resulting from the fastness of fast science entails that scientists could not guarantee the success of their intervention (which is perhaps a general problem for Gardiner's approach, since guaranteeing success of intervention is very demanding). Thus on Gardiner's version of the precautionary principle, many or perhaps all instances of fast science would not warrant precautionary intervention.

The sophisticated version of the precautionary principle defended by Steel ([2015]) suggests a different relation between precautionary reasoning and fast science. Like Trouwborst's version of the precautionary principle ([2006]) and others before it, Steel's is based on a tripod metaphor that differs from the threat tripod metaphor I introduced in §1. Steel's precautionary tripod is based on a knowledge condition, a harm condition, and a recommended intervention. The knowledge condition pertains to how much information we have about a threat, the harm condition pertains to the magnitude of the threat itself, and those in turn recommend a specific intervention which should be aggressive proportional to the plausibility and magnitude of the threat. Like a camera's tripod, the three legs are adjustable, and its balance requires that the adjustment of one leg depends on the adjustments to the others. This allows for useful flexibility in applying the precautionary principle in different kinds of scenarios.

Features of fast science influence the balance of Steel's precautionary tripod. The more that the Principle of Similarity is violated in the detection phase, the less certain would be the estimation of the plausibility, magnitude and imminence of the threat, and thus the knowledge leg of Steel's tripod would be minimised, and so by Steel's proportionality requirement, the recommended intervention should be less aggressive than if the Principle of Similarity were better satisfied. The more that the Principle of Similarity is violated in the intervention phase, the more uncertain we should be about the intervention's effects, and if that uncertainty were great enough, then the very same precautionary principle that favours the intervention could also speak against the intervention (and that would in turn violate another requirement of Steel's precautionary principle, namely the consistency requirement, which holds that an intervention not be recommended against by the same precautionary principle that recommended in favour of it).

Even if a decision procedure—precautionary or not—warrants implementing a particular intervention, that justification should have a temporal restriction, because during the time that the intervention is being implemented better-quality scientific work can be performed, the results of which could then recommend against the intervention (a point argued by Winsberg et al. [2020] and White et al. [2022] in the context of Covid lockdowns).

Some threats are great enough to eliminate all life on Earth, or at least all human life. Bostrom defines an 'existential risk' as a risk that 'threatens the premature extinction of Earth-originating intelligent life or the permanent and drastic destruction of its potential for desirable future development' ([2013], p. 15). Examples include those mentioned above—asteroid impacts, nuclear war, pandemics—and also artificial intelligence, climate change, hostile aliens, and gamma-ray bursts, among many others. If these threats could provide an *in extremis* justification for engaging in and acting on fast science, scientists could flout the reliability-inducing principles and practices of normal science on a large scale, given the number of existential threats and their magnitude. That, in turn, could have harmful epistemic consequences, as our study of these threats and means to intervene on them would become less reliable, and could have harmful practical consequences, as interventions could be costly and failed interventions could be catastrophic. Existential risks clearly involve threats of a great magnitude, and they excite interest because they are, to some degree, plausible. Yet they are less imminent than, say, the Covid-19 pandemic was in the spring of 2020. Probably hundreds of millions of year will pass before a large enough and near enough gamma-ray burst ends life on Earth. One might respond by saying that imminence should be construed not merely in terms of how far off in time a threat is but also in terms of how much time is required to develop an effective intervention against that threat, and if so, then the above existential risks may very well be imminent. Yet, any estimate of how much time it will take to develop an effective intervention against threats like gamma-ray bursts or hostile aliens would be, I believe, a flight of fancy, in part because it is impossible to predict the future surprises of science. Putting this nuance aside, typical existential risks do not warrant fast science. The existential risk of climate change is much more imminent, but still there is time—decades, at least—to perform plenty of routine science before a climate apocalypse ends it all.

6. Conclusion

We have just emerged from a time in which an episode of fast science turned our lives upside down. With a few exceptions, at the time philosophers of science had little to say about it. Our discipline, slow and steady, is accustomed to studying routine science, a tortoise watching tortoises. Yet science is sometimes more like a hare.

I have argued that supreme emergencies can justify engaging in and acting on fast science. If a threat is catastrophic, imminent, and plausible, then scientists can cut corners to quickly recommend interventions. Yet fast science should be as close as possible to routine science in its use of the reliability-inducing principles and practices of routine science, and those principles and practices which are violated when developing an intervention for a supreme emergency should not be the same as those violations which were used to estimate the threat tripod. Fast science cannot be a tortoise, but it ought not be a hare.

Acknowledgements

For written commentary and discussion I am grateful to Ina Jäntgen, Cristian Larroulet Philippi, Benjamin Chin-Yee, Adrià Segarra, Claudio Davini, Eric Winsberg, Michael Shao, Jamie Shaw, Tarun Menon, Rhys Southan, Hayden Wilkinson, Andreas Mogensen, and Nancy Cartwright. Three referees gave incredibly helpful input, as did one of the journal's editors. I am also grateful for discussion with audiences in Dubrovnik, Cologne, Durham, and Bielefeld.

Department of History and Philosophy of Science
University of Cambridge
Cambridge, United Kingdom
jms303@cam.ac.uk

References

Biddle, J. and Kukla, R. (Q.). [2017]: 'The Geography of Epistemic Risk', in K. Elliott and T. Richards (*eds.*) *Exploring Inductive Risk*, New York: Oxford University Press, pp. 215–238.

Birch, J. [2021]: 'Science and Policy in extremis: the UK's Initial Response to Covid-19', *European Journal for Philosophy of Science,* **11**:90.

Bostrom, N. [2013]: 'Existential Risk Prevention as Global Priority', *Global Policy,* **4,** pp. 15–31.

Bradley, R. and Steele, K. [2015]: 'Making Climate Decisions', *Philosophy Compass,* **10/11**, pp. 799–810.

Broome, J. [2012]: *Climate Matters: Ethics in a Warming World,* New York: Norton.

Brown, M. and Havstad, J. [2016]: 'The Disconnect Problem, Scientific Authority, and Climate Policy', *Perspectives on Science,* **25**, pp. 67–94.

Claveau, F. and Grenier, O. [2019]: 'The Variety-of-Evidence Thesis: A Bayesian Exploration of its Surprising Failures', *Synthese,* **196**, pp. 3001–3028.

Douglas, H. [2000]: 'Inductive Risk and Values in Science', *Philosophy of Science* ,**67**, pp. 559–579.

Elliott, K. and McKaughan, D. [2014]: 'Nonepistemic Values and the Multiple Goals of Science', *Philosophy of Science,* **81**, pp. 1-21.

Elliott, K. [2022]: 'A Taxonomy of Transparency in Science', *The Canadian Journal of Philosophy,* **52**, pp. 342–355.

Erasmus, A. [Forthcoming]: 'The Bias Dynamics Model: Correcting for Meta-Biases in Therapeutic Prediction', *Philosophy of Science*

Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., et al. [2020]: 'Report 9: Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand', *Imperial College COVID Response Team. March 16.*

Friedman, D. & Šešelja, D. [2023]: 'Scientific Disagreements, Fast Science and Higher-Order Evidence', *Philosophy of Science,* **90**, pp. 937–957.

Gert, B. [1998]: *Morality: Its Nature and Justification,* New York: Oxford University Press.

Havstad, J. and Brown, M. [2017]: 'Neutrality, Relevance, Prescription, and the IPCC', *Public Affairs Quarterly,* **31**, pp. 303–24.

Heesen, R. [2018]: 'Why the Reward Structure of Science Makes Reproducibility Problems Inevitable', *The Journal of Philosophy,* **115**, pp. 661–674.

Hughes, J. [2003]: *The Manhattan Project: Big Science and the Atomic Bomb*. New York: Columbia University Press.

Jäntgen, I. [Forthcoming]: 'The Dominance of Absolute Outcome Measures', *Philosophy of Science*

Kitcher, P. [1982]: *Abusing Science: The Case Against Creationism,* Boston: MIT Press.

Koskinen, I. [2020]: 'Defending a Risk Account of Scientific Objectivity', *The British Journal for the Philosophy of Science,* **71**, pp. 1187–1207.

Kourany, J. [2010]: *Philosophy of Science After Feminism,* New York: Oxford University Press.

Kuorikoski, J. and Marchionni, C. [2016]: 'Evidential Diversity and the Triangulation of Phenomena', *Philosophy of Science,* **83**, pp. 227–47.

Larroulet Philippi, C. [2022]: 'There is Cause to Randomize', *Philosophy of Science,* **89**, pp. 152–170.

Longino, H. [1990]: *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry,* Princeton: Princeton University Press.

Longino, H. [2004]: 'How Values Can be Good for Science', in P. Machamer, and G. Walters (*eds.*), *Science, Values, and Objectivity,* Pittsburgh: University of Pittsburgh Press, pp. 127–142.

Merton, R. [1942/1973]: 'The Normative Structure of Science', in *The Sociology of Science: Theoretical and Empirical Investigations,* Chicago: University of Chicago Press, pp. 267–278.

Miller, B. [2016]: 'Scientific Consensus and Expert Testimony in Courts: Lessons from the Bendectin Litigation', *Foundations of Science,* **21**, pp. 15-33.

Northcott, R. [2022]: 'Pandemic Modeling, Good and Bad', *Philosophy of Medicine,* **3**.1.

Pielke, R. [2007]: *The Honest Broker: Making Sense of Science in Policy and Politics,* Cambridge: Cambridge University Press.

Posner, R. [2004]: *Catastrophe: Risk and Response,* New York: Oxford University Press.

Rudner, R. [1953]: 'The Scientist Qua Scientist Makes Value Judgements', *Philosophy of Science* **20**, pp. 1–6.

Sandin, P. [2009]: 'Supreme Emergencies Without the Bad Guys', *Philosophia,* **37**, pp. 153–167.

Schroeder, A. [2021]: 'Democratic Values: A Better Foundation for Public Trust in Science', *The British Journal for the Philosophy of Science,* 72, pp. 545–562.

Schupbach, J. [2018]: 'Robustness Analysis as Explanatory Reasoning', *The British Journal for the Philosophy of Science,* **69**, pp. 275-300.
Shaw, J. [2022a]: 'On the Very Idea of Pursuitworthiness', *Studies in History and Philosophy of Science,* **91**, pp. 103–12.
Shaw, J. [2022b]: 'Revisiting the Basic/Applied Science Distinction: The Significance of Urgent Science for Science Funding Policy', *Journal for General Philosophy of Science,* **53**, pp. 477–99.
Sheldrick, K., Meyerowitz-Katz, G. and Tucker-Kellogg, G. [2022]: 'Plausibility of Claimed Covid-19 Vaccine Efficacies by Age: A Simulation Study', *American Journal of Therapeutics,* **29,** pp. e495–e499
Steel, D. [2016]: 'Accepting an Epistemically Inferior Alternative? A Comment on Elliott and McKaughan', *Philosophy of Science,* **83**, pp. 606–612.
Steel, D. [2015]: *Philosophy and the Precautionary Principle: Science, Evidence, and Environmental Policy,* Cambridge: Cambridge University Press.
Steele, K. [2012]: 'The Scientist Qua Policy Advisor Makes Value Judgments', *Philosophy of Science*, **79**, pp. 893–904.
Stengers, I. [2018/2013]: *Another Science is Possible: A Manifesto for Slow Science.* Cambridge: Polity Press.
Sunstein, C. [2005]: 'Cost-Benefit Analysis and the Environment', *Ethics,* **115**, pp. 351–385.
Trouwborst, A. [2006]: *Precautionary Rights and Duties of States.* Leiden: Brill.
van Basshuysen, P. and White, L. [2021]: 'Were Lockdowns Justified? A Return to the Facts and Evidence', *Kennedy Institute of Ethics Journal*, **31**, pp. 405–428.
van Basshuysen, P., White, L., Khosrowi, D., and Frisch, M. [2021]: 'Three Ways in Which Pandemic Models May Perform a Pandemic', *European Journal for Philosophy and Economics,* **14**, pp. 110–127.
Walzer, M. [1973]: 'Political Action: The Problem of Dirty Hands', *Philosophy and Public Affairs,* **2**:, pp. 160–180.
Walzer, M. [2000]: *Just and Unjust Wars: A Moral Argument with Historical Illustrations,* (3rd ed.), New York: Basic Books.
White, L., van Basshuysen, P., and Frisch, M. [2022]: 'When is Lockdown Justified?', *Philosophy of Medicine,* **3**, pp. 1–22.
Wilholt, T. [2013]: 'Epistemic Trust in Science', *The British Journal for the Philosophy of Science,* **64**, pp. 233–53.
Winsberg, E., Brennan, J., and Surprenant, C. [2020]: 'How Government Leaders Violated Their Epistemic Duties During the SARS-CoV-2 Crisis', *Kennedy Institute of Ethics Journal*, **30**, pp. 215–242.
Winsberg, E., Brennan, J., and Surprenant, C. [2021]: 'This Paper Attacks a Strawman but the Strawman Wins: A reply to van Basshuysen and White', *Kennedy Institute of Ethics Journal,* **31**, pp. 429–446.
Winsberg, E. [2022]: 'Who is Responsible for Global Health Inequalities after Covid-19?', *Glob Epidemiol.,* 100081.