# Creativity in the social epistemology of science

Mike D. Schneider[*][†]

**Abstract**

Currie (2019) has introduced a novel account of creativity within the social epistemology of science. The account is intended to capture how conservatism can be detrimental to the health of inquiry within certain scientific communities, given the aims of research there. I argue that recent remarks by Rovelli (2018) put pressure on the applicability of the account. Altogether, it seems we do not yet well understand the relationship between creativity, conservatism, and the health of inquiry in science.

# 1  Introduction

Currie (2019) argues that research in existential risk ('X-risk') should be more creative than it likely is, given the realities of contemporary scientific practice. In the course of the argument, he introduces a general account of creativity in scientific discovery (hereafter, 'creativity'). This account is intended to capture how conservatism can be detrimental to the health of inquiry in scientific communities, given certain aims of research. It is also advertised as complementing the use of formal modeling in studying policy initiatives within the social epistemology of science.

Independent of Currie's project, Rovelli (2018) decries a "why not?" ideology he reports is in vogue within his scientific community, engaged in fundamental physics research. By his reckoning, this ideology promotes a method of guesswork. His concern is that such a method is detrimental, given facts about his community and their research aims.

Here, I will argue that Rovelli's remarks, when interpreted in light of Currie's account, raise trouble for the general applicability of the latter. Evidently, Currie's account fails to countenance the possibility that revolutionary theorizing might be valuable, as features in Rovelli's argument. But since it is difficult to discern when revolutionary theorizing is likely not valuable to a community, it is unclear when Currie's account may be deemed appropriate for studying the effects of conservatism on the health of inquiry therein. This threatens to undermine the use of such an account in arguments undergirding policies meant to respond to conservatism. It would be prudent to seek out means of identifying what it is about any given scientific community that could render Currie's account appropriate there.

# 2 Creativity in science

Stanford (2019) has argued that the structures and institutions of contemporary science foster conservatism in research, stifling revolutionary theorizing. Currie (2019) is concerned that the same conservatism is detrimental to inquiry within X-risk. This is because, according to Currie, disciplines like X-risk are best pursued creatively. Arguing that creativity is in tension with conservatism, Currie concludes that the scientific communities focused on disciplines like X-risk are likely insufficiently creative— the structures and institutions of science stack the deck against the disciplines' prospects.

As just presented, Currie's project depends essentially on his providing an explicit account of creativity within a scientific community. The remainder of this section is dedicated to describing the account he provides, as well as developing it further (where necessary) in a friendly manner.

Consider the situation wherein there is some well-posed problem, whose solution a scientific community agrees constitutes the aim of their collective research. The statement of the problem places severe constraints on what counts as viable research within that community, united by that aim. We may think of the statement of the problem as characterizing the research program pursued by that community. And associated with that problem is, following Currie, a collection of possible solutions. This 'solution space' is meant to be roughly coextensive with all professional moves available to members of that community, engaged in that research program. The researchers occupy points in the solution space, and they choose which points to occupy next.[1]

---

[1] In fact, there are other professional strategies that are ultimately available to researchers, regarded as decision-making agents. Whether activity gets channeled into those other strategies, rather than into moving between solutions, is an important degree of freedom in Currie's account.

Currie introduces into this picture the following two metaphors. 'Hot searches' through solution space are energetic; 'cold searches' are the opposite. A hot search refers to a sequence of points, whose iterative selection by a theorist describes that theorist as hopping around through the solution space. A cold search refers to a similar sequence of points, except that it describes the behavior of a theorist who is nearly staying still.

To make these metaphors, Currie needs a notion of distance between points in the space. He borrows from Bayesian epistemology to develop one. (I will have more to say that is critical of this below.) By his reckoning, distances to solutions are relativized to each individual at a time, and are indexed to that individual's credences at that time. So, roughly speaking, solutions assigned low priors are far, and solutions assigned high priors are near.[2]

Currie does not elaborate on the interpretation of these priors. Evidently, he has in mind something pragmatic: "Our priors serve to set expectations across a space of possible solutions to a problem" [p. 6]. In this respect, the account is non-committal about what it is that ultimately makes a solution worth visiting. We are free to suppose that there is some unspecified constellation of virtues, possibly specific to the research program at hand, that one hopes is jointly maximized (i.e. via some method of aggregation) by whatever solution is visited next. On this picture, hot searches are sequences for which the researcher's decisions are insensitive to their beliefs about where it will be prudent to visit. Oppositely, cold searches occur when the researcher's choices correlate strongly with those beliefs.

Currie then defines an agent's creativity in terms of their propensity for hot searches.

---

[2]As will become clear, it may be that we ought to insert a ceteris paribus clause here. If so, we would say that whatever are otherwise the distances to solutions, those numerical values are then systematically deformed to reflect comparative facts about one's priors over each solution.

In other words, an agent is creative in proportion to the unconditional probability that they attempt a distant, low-credence solution. A community's creativity, meanwhile, is defined to correspond with what would generally occur if the members of the community were all individually creative. The upshot is that a community's creativity is defined as proportional to the efficiency with which they explore solution space widely. (What it means to explore widely is, of course, agent-relative. Here, we might assume that a community explores widely when it does so by the lights of most of its members.)

This wide exploration of solution space is in contrast with what, following Currie, we may call 'pooling'. Intuitively, pooling occurs when individuals within the community fail to be creative, each favoring cold searches instead of hot searches. But, as Currie notes, pooling may be avoided in such a case, provided that the community is cognitively diverse. So long as cognitive diversity is understood in terms of diverse distributions of priors, cognitively diverse individuals engaging in cold searches will, collectively, explore widely. This community would count as creative, according to Currie, even though the individuals who comprise it do not.

The creativity of a community is therefore not uniquely determined by facts about the creativity of its constituents. Their propensity for peer disagreement (and so, the social structure of science, etc.) also matters. And according to this view, a community may be made more creative in various ways. One way is by interventions to promote sustained cognitive diversity, as we have understood it here. Another is by incentivizing hot searches, or increasing creativity at the individual level. In both cases, pooling is reduced, in favor of wider exploration.[3]

---

[3]A third way to increase creativity, noted by Currie, is to impose on the community a diverse collection of search algorithms. But this raises a question: what distinguishes, in practice, our imposing a diverse collection of search algorithms from our incentivizing hot searches? At the

Building on recent work by Stanford (2019), such interventions are, according to Currie, in contrast with the unchecked effects of conservatism in professional science today. This is because, according to Currie, conservatism promotes pooling, as we have understood it here. But depending on the given research program, it may or may not be detrimental that science today is, generally, conservative. This is because a research program ought to be assessed individually, according to the "local details" [p. 3] relevant to it. Those details determine, for instance, whether the community is better off investing in strategies other than those relevant to scientific discovery (cf. footnote 1 above). If so, any resulting pooling according to shared priors need not be unhealthy.

As just stated, the utility of Currie's account is ultimately going to rest on certain further facts: which kinds of local details ought we to recognize as rendering creativity— as opposed to pooling— a standard of good epistemic health in the community? Such local details are encoded, we may suppose, in the statement of the problem that constitutes the aim of that community's research. Recall that it is from this problem that, in principle, we may extract the parameters of the solution space we envision the community to explore. It follows that assessments of the local details of a research program will generally shape our expectations about the solution space associated with the problem. Likewise, facts about a solution space can correlate with facts about whether pooling or creativity is preferred in the corresponding research program.

Unfortunately, Currie does not state how such a correlation would work. This omission could suggest that we ought not to regard local details as shaping our

---

level of analysis presently provided, it is unclear that there is any distinction. As suggested in footnote 2, it may be that we should ultimately think of solution space as admitting some intrinsic structure, independent of credences. In that case, search algorithms could be defined with respect to that intrinsic structure, and would generally result in searches that appear hot.

expectations about solution space (besides via shaping our priors). But this would render Currie's account in tension with the standard interpretation of formal landscape models. Currie regards the use of such models within the social epistemology of science as complementing his approach (cf. p. 11 in the article). In such models, one typically regards the intrinsic structure of the landscape as an independent variable, whose possible values encode arbitrary research environments. So too, we might conclude, the structure of a solution space should reflect facts about the corresponding research program.

In light of this, I think it is appropriate to regard Currie's discussion of X-risk as illustrating the reasoning that would shape the relevant solution space. His ultimate conclusion is that X-risk should be creative because it should be "multi-disciplinary, pluralistic, and opportunistic" [p. 26]. We might speculate, on the basis of this, that the local details relevant to the problem of X-risk render the solution space as unusually vast.[4] In a vast solution space, cold searches could seem unfruitful, no matter how cognitively diverse we may plausibly imagine are the researchers. Consequently, creativity is generally preferred in such a case, consistent with Currie's reasoning about X-risk.

To recap: treating research programs as solution spaces, creativity is a matter of how the relevant communities explore those spaces, given priors. Conservatism encourages pooling according to shared priors, which is opposite creative exploration. But specific facts about the solution space at hand can determine, in a given community, which of

---

[4]There is room for disagreement here. For instance, Currie's discussion of X-risk places some emphasis on its normative aspect— i.e. threat mitigation— and its role in the public eye. It is not clear what these would have to do with the size of the solution space. This ambiguity motivates a revisionist attitude toward distances in the space. (See also footnotes 2 and 3 above.)

creativity or pooling is likely preferred. Those facts are ultimately grounded in the statement of the problem identified by that community as constituting their research program.

# 3   The situation in fundamental physics

Consider now the article by Rovelli (2018). Rovelli is a theoretical physicist focused on quantum gravity, the problem that characterizes fundamental physics research today.[5] Indeed, we may understand the problem of quantum gravity to be that which shapes the relevant solution space, against which creativity in fundamental physics is to be assessed. In what follows, I take Rovelli to have expertise regarding that solution space, as well as privileged access to it.

Rovelli's article is adversarial. Our attention is best directed to a passage that comes in the middle, immediately following his presentation of what he calls the "why not?" ideology. According to Rovelli, this uncritical ideology is responsible for the rise of a damaging method of guesswork in contemporary fundamental physics practice. According to the method, reason need not be (nor can be, fruitfully) given to merit the study of any new research proposal. The criticism of the method proceeds as follows [p. 7]:

> Arbitrary jumps in the unbounded space of possibilities have never been an effective way to do science. The reason is twofold: first, there are too many possibilities, and the probability of stumbling on a good one by pure

---

[5]This is, of course, a massive simplification. But so too is the problem characterizing X-risk in Currie's project. Whether the simplification is tolerable despite such objections depends on the particular context of its use.

chance is negligible; but more importantly, nature always surprises us and we, the limited critters that we are, are far less creative and imaginative than we may think. When we consider ourselves to be "speculating widely", we are mostly playing out rearrangements of old tunes: true novelty that works is not something we can just find by guesswork.

As in Currie's article, we have here a spatial account of scientific discovery. Scientists decide how to move amongst points in the space (now, of 'possibilities', rather than 'solutions'). The role of the "why not?" ideology is to support a method of guesswork. We can understand this method as a decision procedure, the repeated execution of which amounts to "arbitrary jumps" in the space. (More formally, we might think of such a method as analogous to Monte Carlo sampling, with respect to some unspecified probability distribution on the space. Based on the context surrounding the quoted passage, Rovelli clearly has in mind a distribution that is meant to be uncorrelated with one's priors.) But absent any greater detail about the account Rovelli envisages, it is unclear why such a method should be as damaging as he claims. Prima facie, Currie's account of creativity should be helpful as a means to interpret the argument.

In Currie's framework, Rovelli's 'space of possibilities' may be understood as a solution space for the problem of quantum gravity. The solutions to the problem are, then, candidates for what may turn out to be a satisfying theory of quantum gravity. Given this reading, Rovelli's principal claim about the space is that it is vast. This seems right. In other contexts, this space is taken to be synonymous with 'theory space', the collection of all possible fundamental theories (see, e.g. (Dardashti, 2019)). From here onward, I will adopt this 'theory space' language when talking about the space of

solutions relevant to the problem of quantum gravity.[6]

Recall that creativity at the community level is spelled out, on Currie's account, in terms of exploring widely in the relevant solution space. I have suggested that we understand Rovelli's remarks in terms of fundamental physicists exploring the vast theory space corresponding to the problem of quantum gravity. Since the space is vast, by the argument at the end of the previous section, creativity is likely preferred to pooling. In other words, a more creative community is likely better off, given the local details of the problem of quantum gravity. Wider exploration should be good here.

Meanwhile, fundamental physicists are, according to Rovelli, *uncreative* (or, at least, are "far less creative" than they may think).[7] On the present interpretation, this would suggest that fundamental physicists fail to explore widely. Increasing creativity should be desirable.

Naively, guesswork is one such method to do so. (As described above, except if the sampling is with respect to a probability distribution correlated with one's priors, guesswork will generally produce hot searches.) On Currie's account, we may thereby understand Rovelli to hold the view that the method of guesswork happens to be implemented poorly by his community. Moreover, according to Rovelli, when his community engages in guesswork, they fail to speculate as "widely" as they typically believe themselves to speculate. So: the community does not explore widely, and they fail to recognize that this is the case.

This seems to provide a sufficient reason that the method is, according to Rovelli,

---

[6]In (Schneider, 2020), I criticize the relevance of this 'theory space' view in assessing the methodology of quantum gravity research.

[7]What relation this testimony could bear to the broader conversation about conservatism in science is interesting to consider, but a tangent at present.

damaging. Because theory space is vast, creativity constitutes a standard of good epistemic health in contemporary fundamental physics. Meanwhile, the community's poor implementation of guesswork fosters an exaggerated perspective as to how healthy their inquiry really is. Our initial hunch was correct: Currie's account of creativity can help us get traction on Rovelli's argument.

Yet, there is something unsatisfying about this interpretation of the argument. Consider the reason that Rovelli supplies for his testimony that the community implements the method of guesswork poorly. The poor implementation is due to the fact that "we, limited critters that we are, are far less creative and imaginative than we may think". In other words, guesswork is implemented poorly by his community, because their being limited ensures that they *cannot implement it well*. In particular, it his community's lacking creativity (and imagination), on this interpretation, that ultimately bears responsibility for the method being damaging.

Whether Rovelli's argument is compelling, so interpreted, is therefore going to turn on whether a community's lacking creativity can be understood to intervene on the efficacy of a method they attempt to employ. And here, Currie's account provides little guidance. Facts about the community's pooling with respect to shared priors cannot obviously prohibit researchers, all of whom are willing to speculate irrespective of their priors, from doing so. In this respect, Rovelli's argument depends on creativity (or the lack thereof) playing a further role in the social epistemology of his community than is readily countenanced by Currie's account.

Note that this observation does not present an objection to Currie's argument, as his argument does not require that his account of creativity be complete. Nonetheless, as I will now discuss, Rovelli's argument is ultimately compelling, provided that we attribute

11

to Rovelli the view that revolutionary theorizing is valuable in contemporary fundamental physics. And recognizing the importance of such a view to Rovelli's argument should make us wary about assertions that Currie's account is applicable in any particular epistemic situation. Currie's account cannot merely be assumed to capture how to assess the epistemic impact of conservatism on a research program, for which creativity is healthy. A further question about whether or not revolutionary theorizing is valuable complicates the assessment.

# 4    Revolutionary theorizing and the health of inquiry

Suppose that there exist possibilities in theory space that are assigned prior probabilities of zero by all members of the community. Whereas many possibilities are *accessible* to the community, in virtue of being assigned non-zero priors by someone, these further possibilities are *inaccessible*. On Currie's terms, these are possibilities that are located an infinite distance away from the community, and are regarded as infinitely less promising to visit than any accessible possibility.[8]

In such a case, no matter how creative the community is regarding the accessible possibilities, some of theory space will never be explored. So, provided that guesswork fails to be defined over inaccessible possibilities, the method could fail to spread the community as wide as might, ultimately, be desired. This idealized setup sounds

---

[8]Assignments of zero-probability priors to non-contradictions are antithetical to an orthodox Bayesian epistemology. So, it is not obvious that the present supposition, in the case of theory space, is faithful to Currie's project. Nonetheless, given some other structure to the space (cf. footnotes 2-4), we may understand zero-probability priors as an idealization that "pushes off to infinity" the corresponding possibilities. They are, in effect, disconnected from the accessible ones. No amount of information gleaned from work on the latter could ever reign them in.

promising as a means to recover why, according to Rovelli, his community cannot implement guesswork well. We need only to attribute to Rovelli two further claims. The first is that his community's lack of creativity results in there being some possibilities that are inaccessible. The second is that at least some of those inaccessible possibilities are important to the aims of his community's research.

Evidence that Rovelli would endorse each of these claims may be found within the passage already quoted. Namely, what is inadequate about guesswork, says Rovelli, is that it does not yield "true novelty that works". This is because employing it results (instead) in "playing out rearrangements of old tunes". If we interpret the rearrangements of old tunes as the accessible possibilities, his claim is this: what there is to be sought in fundamental physics— i.e. true novelty that works— in fact resides in the inaccessible part of theory space.

Suppose that this reading is correct, and what there is to be sought in fundamental physics is, according to Rovelli, presently inaccessible. Then it is a symptom of the community's not being creative, according to Rovelli, that the implementation of guesswork necessarily fails to engender wide enough exploration. This is because the relevant sampling procedures fail to be defined over the whole of what is *worth exploring*.

We have thereby found a means to articulate the lingering part of Rovelli's argument, which we were unable to do in the previous section. Namely, says Rovelli: what is worth exploring fails to be coextensive with the accessible part of theory space. As a result, guesswork is ineffectual. Worse, employing the method misleads the community in their self-assessment of whether they are sufficiently creative, consonant with their research aims. This is because the method only promotes wide exploration of a kind that is unsuitable for assessing the health of inquiry in fundamental physics. It only

countenances that which is *conceived as* worth exploring (i.e. rather than what *is*).

If this is how we are to understand Rovelli's argument, it is easy to generalize the lesson. Consider any context wherein one has reason to regard the accessible part of solution space as failing to include some of what is worth exploring (putting off, at least for another few paragraphs, the issue of what it means for something to be worth exploring). This is a context in which genuinely revolutionary theorizing is needed, which renders accessible more of the space. In other words, if a community has reason to value revolutionary theorizing in their research, no amount of hot searching amidst that which is conceivable will amount to healthy inquiry. This is despite creativity remaining a standard of good health in that community, given their research aims.

But such a conclusion spells trouble for the applicability of Currie's account in arguments about policy. Currie's observation, as discussed above, is that conservatism promotes pooling with respect to shared priors. To the extent that creativity is anticorrelated with such pooling, Currie concludes research programs that ought to be creative likely suffer, in virtue of conservatism. Therefore, interventions that would promote creativity in the relevant communities would be well motivated, given the broader context of science today. (Indeed, this is just what Currie calls for in the case of X-risk.)

But now, there is cause to doubt that creativity has anything to do systematically with pooling, as defined with respect to shared priors. Creativity may, for instance, be anticorrelated with an entirely different kind of failure to explore, measured against an entirely different distance measure on the space. At least when revolutionary theorizing is valued, this seems to be the case. Indeed, one might even imagine situations wherein pooling, as measured against priors, provides explicit means of playing with what it is

14

that we conceive as worth exploring. (Rovelli seems to have something like this in mind in his advocating for a method built on continuity, in order to break away from playing rearrangements of old tunes.)

If so, interventions to promote creativity cannot be motivated against a background of conservatism, at least as Currie has presented the topic. In cases such as these, we require a different sort of reason to motivate interventions in response to conservatism (when, still, creativity is important). For instance, suppose that the conclusion is warranted: conservatism deprives the relevant community of access to much of solution space (cf. footnote 7). Then it is plausible that what is sought by the community is inaccessible, in which case revolutionary theorizing might be valuable. Policies intended to promote creativity in that community could then be motivated, given the broader conservatism of science today. (And enacting such policies would be all the more important if, following Stanford, we further regard conservatism as stifling revolutionary theorizing.)

On the other hand, we might imagine some cases (perhaps that of X-risk) in which Currie's account adequately captures the effects of conservatism on inquiry. These are cases where we regard a community's capacity for revolutionary theorizing as, antecedently, unimportant to assessing the health of inquiry therein.

Such cases may arise in practice. But if they do, it is very difficult— if not impossible— to reliably identify them as such. What is up for grabs here is our epistemic access to whether that which we presently conceive as worth exploring happens to be coincident with that which is worth exploring. This is one lesson of Stanford's original project, which foremost concerned our means of evaluating the contemporary threats posed by the problem of unconceived alternatives. The upshot is that there may turn out

to be no problem inherent in the applicability of Currie's account in certain cases. Yet, there is a severe problem in asserting when we are reliably in such a case. This matters for the argumentative force of any call for new incentives to promote creativity in any particular community, based on his account. Namely, one must commit to the belief that, whatever it means for a solution to be worth exploring— i.e. given the ultimate aims of the community's research, the individuals' understandings of the problem that shapes that research, and so on— that solution is presently conceived as such.

Whether Currie's account can provide insight into the effect of conservatism on inquiry will therefore require a more sophisticated understanding of creativity. Such an understanding would need to provide a reliable means of picking out those situations wherein the benefits of creativity are not to do with revolutionary theorizing. In those situations, Currie's account could give us some grasp of how to evaluate the epistemic health of the relevant community. But the grounds for that evaluation would ultimately reside in the more sophisticated account. This is because only according to that more sophisticated account could we explain in virtue of what revolutionary theorizing is, in the particular case at hand, rendered unimportant.

# 5 Conclusion

I have argued that Rovelli's remarks ultimately uncover a shortcoming of Currie's account of creativity. This shortcoming concerns the possible value of revolutionary theorizing to the aims of a research program. Lacking a more sophisticated account of creativity, it is difficult to assess a variety of claims of independent interest. For instance, what commitments does Rovelli make about the problem of quantum gravity, in order to

claim that revolutionary theorizing is valuable within contemporary fundamental physics? And when is it appropriate to focus questions about creativity exclusively on just what is conceived as worth exploring? After all, Currie is unequivocal about the relevance of his more narrow account of creativity in the case of X-risk. He states: "...it is this kind of creativity which scientific study of existential risk requires" [p. 8]. So, by what reasons do the local details of X-risk entitle us to restrict our study to an account that disregards the possibility that revolutionary theorizing matters?

Currie anticipates the possibility that a more sophisticated notion of creativity might ultimately be demanded. By his reckoning, this is because his account does not capture 'ingenuity' (p. 8), failing to distinguish creative searches from chaotic ones. Currie then suggests that a new account of creativity, built on the notion of creative 'flair' developed by Gaut (2010), might capture such a distinction.

This suggestion strikes me as promising. For instance, creative searches might be those hot searches that enable the community to subsequently achieve novelty in research (e.g. at the end of some iterative process). But I would like to conclude by noting one major obstruction to developing the suggestion further. Following Currie, the first step in articulating an account of creativity would be to specify how to extrapolate from the individual to the community level. Such a move is essential to an understanding of the relationship between the social structure of science and creativity, like we have understood it here. (Of particular interest is whether conservatism can be responsible for reliably depriving us of access to much of a solution space, within the developed account.) But extrapolating from the individual to the community level is no small challenge. Creative flair is an irreducibly agential notion, concerning an individual's familiarity with their own goals. It is unclear at present what would mark a

17

community that, as a whole, is creative in this refined, goal-sensitive respect.

There is, it seems, still much work to be done.

# References

Currie, A. (2019). Existential risk, creativity & well-adapted science. *Studies in History and Philosophy of Science Part A 76*, 39–48.

Dardashti, R. (2019). Physics without experiments? In *Why Trust a Theory?: Epistemology of Fundamental Physics*, pp. 154–172. Cambridge University Press.

Gaut, B. (2010). The philosophy of creativity. *Philosophy Compass 5*(12), 1034–1046.

Rovelli, C. (2018). Physics needs philosophy. Philosophy needs physics. *Foundations of Physics 48*(5), 481–491.

Schneider, M. D. (2020). What's the problem with the cosmological constant? *Philosophy of Science 87*(1), 1–20.

Stanford, P. K. (2019). Unconceived alternatives and conservatism in science: the impact of professionalization, peer-review, and Big Science. *Synthese 196*(10), 3915–3932.