

Tracking, Reliabilism, and Possible Worlds

Wesley Cooper

Abstract

Robert Nozick's tracking account of knowledge is defended against Colin McGinn's criticisms by drawing on David Deutsch's 'multiverse' conception of possible worlds. Knowledge on the tracking account requires a 'method' or 'way' of believing. Exploiting this feature undercuts the apparent force of McGinn's counter-examples.

This essay focuses on Colin McGinn's critique in *Knowledge and Reality* of Robert Nozick's tracking account of knowledge in *Philosophical Explanations*. Nozick offers a crisp portable statement of the tracking theory as follows, where S is some knowing subject and p is some proposition known.

1. p is true.
2. S believes that p .
3. $\text{not-}p \longrightarrow \text{not-(S believes that } p \text{)}$. (the variation condition)
4. $p \longrightarrow \text{S believes that } p$. (the adherence condition)

The arrow relates antecedent to consequent in the manner of the subjunctive conditional: if the antecedent weren't true, then the consequent wouldn't be true. If Hermia knows that Lysander is waiting for her in the woods, then if Lysander weren't waiting for her in the woods, she would not believe so (variation); and if Lysander were waiting for her in the woods, Hermia would believe that (adherence).

It is important that a more technical and complete statement – about S knowing *via method (or way of believing) M* that *p* – lies in the background.

1. *p* is true.
2. S believes, via method or way of coming to believe *M*, that *p*.
3. If *p* weren't true and S were to use *M* to arrive at a belief whether (or not) *p*, then S wouldn't believe, via *M*, that *p*.
4. If *p* were true and S were to use *M* to arrive at a belief whether (or not *p*, then S would believe, via *M*, that *p*.

The more complete statement is important for putting McGinn's counter-examples to the tracking theory into perspective, and for evaluating his diagnosis that it suffers from being a "local" version of reliabilism as opposed to his own "global" version.

This diagnosis is replaced here by the interpretive frame that McGinn and Nozick offer versions of reliabilism differing about whether capacities or subjunctives should be fundamental, and that when this issue is sorted the two versions play complementary roles in a fuller reliabilist theory, in which neither capacities nor subjunctives is fundamental. McGinn makes a good case that discriminative capacities are significant for the theory of knowledge, but a certain realism about possible worlds shows how the tracking theory might be more basic: capacities imply subjunctive conditionals, and possible worlds might be the truth-makers for these. The many-worlds or multiverse hypothesis is suggested as an appropriate form of realism about possible worlds,

anchoring the relevant counterfactuals in the multiverse, representing all physically possible worlds, while bypassing issues about modal realism pertaining to worlds that are alleged to be knowable *a priori*. The multiverse is knowable *a posteriori* as implied by a scientific hypothesis, not *a priori* as implied by the truth conditions of modal sentences. McGinn is plausibly critical of modal realism of the latter sort, with special reference to David Lewis; but the multiverse version of modal realism is motivated by wholly different considerations than Lewis's, empirical arguments from physics that are untouched by McGinn's criticisms.

Talk about capacities, abilities, powers, dispositions and the like is vague at best. Even when identified with their "categorical bases" in micro-structure, that structure guarantees the capacity only if it supports appropriate subjunctive conditionals. (Opium's "dormitive power" may be such-and-such micro-structure, but only if that structure would put one to sleep if ingested.) Those conditionals in turn backstop the structure by referring to what it actually does in other physically possible worlds. (There is a world in the multiverse where this opium puts one to sleep, though in this world it remains in the medicine cabinet.) The upshot is that McGinn's and Nozick's versions of reliabilism are in effect the same analysis at different levels – less and more fundamental ones, respectively.

McGinn's first counter-example to tracking asks the reader to

[s]uppose we are living in a universe in which there also exists a benevolent deity who watches over our sensory input: he has the intention to preserve this input by artificial means in the event of a cataclysm in which the material objects that actually produce it should suddenly go out of existence. Let us suppose that this cataclysm is, in fact, physically possible and that the deity has the power to carry out his intention. Then it seems that we have the truth of this counterfactual: 'If the objects around me were to go out of existence, I would still believe that I was surrounded by those objects' – since the deity would see to it that my experience sustained this belief were the cataclysm to occur. (We also, of course, have the adherence condition satisfied in this case.) Yet I am reluctant to say that, because of these facts, we do not *know* that we are surrounded by material objects: for the truth of the counterfactual does not, intuitively, make our true belief that we are surrounded by material objects merely *accidentally* true. Suppose that in the whole history of the universe the cataclysm never in fact occurs, though *if* it had the deity would have intervened to preserve our beliefs: can we really say that we do not then know, for example, that the earth exists? (McGinn 1999, pp. 8-9)

But consider two ways of knowing that the denizens of this universe might employ, an immanent method that relies exclusively on sense experience, and an immanent -plus or transcendental method that is proof against deception by transcendent entities. Disclaiming a transcendental method allows the denizens to know. The immanent method might be restated as requiring, as McGinn writes, “some condition that speaks of the person’s propensity to believe the truth with respect to a range of distinct ‘relevant’ propositions.” Insistence on this requirement is the distinctive feature of his global reliability theory. Since this is to be incorporated into the tracking account, it is fair to specify the method in this way. There is no reason to polarize McGinn’s reliabilism as global and Nozick’s as local, because Nozick’s analysis includes reference to method as well as the four conditions in the portable version, and method can be specified, as it just was, so as to avoid counter-examples. McGinn attends to the portable statement of the

tracking theory, dismissing the references to method in the fuller statement as “some minor refinements.” This neglect is responsible for the polarizing.

In passing, note that McGinn’s assumption about the physical possibility of the benevolent-deity-universe is not innocent, at least when the tracking theory is integrated with the multiverse hypothesis. True counterfactuals are grounded in actual events in the multiverse, so the ones that figure in the counter-example are false, as surely as the physics-defying antics in Road Runner cartoons.

Having defended the necessity of the variation condition, turn now to McGinn’s counterexample to its sufficiency.

You visit a hitherto unexplored country in which the inhabitants have the custom of simulating being in pain. You do not know that their pain behaviour is mere pretence, and so you form the belief of each person you meet that he or she is in pain; imagine you have acquired a great many false beliefs in this way. There is, however, one person in this country who is an exception to the custom of pain pretence: this hapless individual *is* in constant pain and shows it (we can suppose that he falsely believes others to be in his unfortunate condition – he has not been told of the pretence by the others). You also believe of this person, call him *N*, that he is in pain. Now I take it that we would not say that your true belief that *N* is in pain counts as knowledge, for it is, intuitively, a mere accident that your belief is true in this instance. But now consider the relevant counterfactuals, in particular ‘if *N* were not in pain, you would not believe that *N* was in pain’: this counterfactual is *true* in the envisaged circumstances, since if *N* were not in pain then (unlike the pretenders around him) he would not behave as if he was, and so you would not believe that he was. So your belief that *N* is in pain *does* track the truth of that proposition even though it does not rank as knowledge. (McGinn 1999, pp, 11-12)

But consider two methods or ways of knowing that you might be using. One relies on induction from the pretenders to the hapless individual. The other is a one-off method: it sizes up the individual as a pain-sufferer without the benefit of inference from observation of the pretenders. If you were using the one-off method, your belief about the individual would be knowledge. But the counter-example stipulates in effect that you are using the other method, and consequently your belief is tainted – not by the insufficiency of the variation condition, but by the method you used, which doesn't mesh well with this hitherto unexplored country. The method includes a folk-psychological theory, in particular about when pain behavior can be expected, that happens to be wildly misleading in this country.

These replies to the counterexamples suggest that the differences between McGinn's and Nozick's versions of reliabilism are merely notational. The references to a broad range of propositions in McGinn are brought into Nozick by a "method or way of knowing." Moreover, McGinn requires local reference to a specific known proposition, like the one that figures in the variation and adherence conditions, in order to cope with counterexamples similar to the ones he poses. So local-versus-global doesn't frame helpfully the relationship between the two versions of reliabilism.

Suppose again that we are living in the universe of the benevolent deity. This time however a cataclysm actually occurs at noon on a certain day, and the deity implements

his policy of intervening to paper it over, producing in us convincing but illusory experience as of material objects. Before noon and after noon there really are material objects, and correspondingly we have knowledge of a range of propositions about them. But at noon we lack such knowledge, and the corresponding proposition about our knowledge of them is false. So McGinn's analysis of *S knows that p* must include not only reference to a range of propositions about which S can discriminate truth from falsehood, such as those about the world before and after noon, but also explicit provision that *p* belongs in that range (or not, as in the preceding revision of McGinn's thought-experiment). Like Nozick's analysis, it must have local features as well as global ones. This can be brought out as well in the thought-experiment about the unexplored territory, where the traveller might figure out the widespread pretence and notice something about the genuine pain sufferer that sets him apart. His knowledge of this person's pain is not tainted by the widespread pretence, as it is in the first version of the experiment. McGinn's analysis of *S knows that N is in pain* must do the local work of determining whether S is tracking N's being in pain, in addition to the global work of specifying whether S's method of knowing is tainted by widespread pretence.

McGinn would not want to rephrase his analysis in these terms, preferring brute appeal to discriminative capacities over the tracking theory's appeal to subjunctives. He writes:

The point is that it is unsatisfactory to employ counterfactuals in a primitive way in one's analysis of categorical propositions; they have *dependent* truth value. We can always legitimately ask what *makes* a given counterfactual true and expect to be presented with a suitable categorical fact. Now it seems to be that this general thesis imposes a constraint upon philosophical analyses, to the effect that we should be able to say what categorical propositions ground the counterfactuals we employ in the analysis....if non-circular categorical grounds *can* be produced it seems that the counterfactuals are in principle dispensable in the analysis; they serve merely as an eliminable intermediate or interim step to the real analysis, which is categorical in form....[W]e are entitled to press Nozick on the question what makes his tracking counterfactuals true: what *categorical* facts about the believer S and S's relation to the world *make* it true that if it weren't the case that *p* S would not believe that *p* and if it were S would? (McGinn 1999, p. 16)

He believes that a satisfactory analysis would reveal the categorical facts upon which subjunctive conditionals depend. This is not unreasonable, but on the other hand his own preference for capacities is questionable for the reasons given earlier, about vagueness and inseparability from subjunctives. Is there a deeper level of analysis?

Both Nozick and McGinn are willing to rephrase the tracking theory in terms of possible worlds. Nozick writes,

This point [about the power and intuitiveness of the subjunctive condition] is brought out especially clearly in recent 'possible-worlds' accounts of subjunctives: the subjunctive is true when (roughly) in all those worlds in which *p* holds true that are closest to the actual world, *q* also is true. (Examine those worlds in which *p* holds true closest to the actual world, and see if *q* holds true in all these.) Whether or not *q* is true in *p* worlds that are still farther away from the actual world is irrelevant to the truth of the subjunctive. (Nozick 81, p. 174)

And in the same vein McGinn writes,

Putting these two conditions [variation and adherence] in terms of the usual possible worlds semantics for counterfactuals, we can say that S has knowledge that p iff (i) in all possible worlds closest to the actual world in which p does not hold S does not believe that p , and (ii) in all those close worlds in which p does hold S believes that p : belief that p is not preserved in the close not- p worlds and it is preserved in the close p worlds. (McGinn 1999, p. 9)

A possible-worlds account of subjunctives/counterfactuals is formal, in that it does not entail or exclude the various conceivable interpretations that give it content. Not excluded in particular is an interpretation that construes its possible worlds as the physically possible worlds of the multiverse. Lewisian modal realism is another interpretation, one that isn't restricted to physically possible worlds and, at least in Lewis's favoured account, does not allow interaction between worlds (though he explores the logic of overlapping worlds). That multiverse worlds aren't the same as Lewis's doesn't prevent the former from belonging to a legitimate interpretation of possible-worlds semantics. Nozick and McGinn regard this semantics as nothing more than a paraphrasing or formalizing device, but their agreement on this point does not settle the matter about whether the multiverse hypothesis can interpret its worlds more deeply, as grounding subjunctives. Nozick disclaims anything more (p. 81), and McGinn argues at length against Lewisian modal realism, which attempts to extract ontological consequences from the truth conditions of modal sentences. Although he surmises that "Lewis's metaphysics is the only way to make clear and honest sense of an ontology of possible worlds," the horizon of this claim is limited to theorists like Lewis, Robert Stalnaker, and Saul Kripke who derive the ontology *a priori*. There is a clear path therefore for a narrow, *a posteriori*

derivation of a possible-worlds ontology that grounds subjunctives. Nozick's disclaimers, Lewis's defense of modal realism, and McGinn's skepticism about it leave that path open. (I am indebted in this paragraph to Jim Stenberg.)

David Deutsch's account in *The Fabric of Reality* serves as an accessible statement of the multiverse interpretation of these experiments. (See also Deutsch 1985.) They have to do with interference phenomena, which on the multiverse account are the effects of a particle in one universe on its counterpart in another. So when a photon is shot through a slit toward a barrier in an interference experiment, the resulting pattern on the barrier is *as if* the photon had collided with an invisible "shadow" photon coming through one of the other slits. The multiverse hypothesis takes this appearance at face value. Other options include the hypothesis that the photon exhibits wave-particle duality (Bohm's interpretation), and the hypothesis that science should refrain from metaphysical speculation about what's really going on in such experiments (the Copenhagen interpretation). Deutsch makes a case for scientific realism over instrumentalism and consequently rejects the Copenhagen interpretation and its like. He suggests that the issue between Bohm's allegedly single-universe interpretation of quantum theory and the many-worlds view is ultimately to be decided by appeal to complexity theory.

Deutsch assumes that degree of complexity will be a function of amount of computation, reasoning that

[w]orking out what Bohm's invisible wave will do requires the same computations as working out what trillions of shadow photons will do. Some parts of the wave describe us, the observers, detecting and reacting to the photons; other parts of the wave describe other versions of us, reacting to photons in different positions. Bohm's modest nomenclature – referring to most of reality as a 'wave' – does not change the fact that in his theory reality consists of large sets of complex entities, each of which can perceive other entities in its own set, but can only indirectly perceive entities in other sets. These sets of entities are, in other words, parallel universes. (Deutsch 1997, p. 56)

So Bohm's variables are in effect under-interpreted and not fully explanatory. Just as Ptolemaic epicycles, if fully interpreted/explanatory, would give a Galilean description of the movement of planets, so too Bohmian variables would yield a multiverse. The epicycles are really tracking Galilean motion, and the variables are really tracking multiverse phenomena. (For a philosopher's defense of Bohm see Christopher Norris's *Quantum Theory and The Flight from Realism*. See that work and also his "Should philosophers take lessons from Quantum Theory" for criticism of Deutsch. For a scientific journalist's sympathetic discussion of Deutsch, see Julian Brown's *The Quest for the Quantum Computer*.)

Deutsch holds that classical spacetime physics, understood deterministically as it is on his view, implies that counterfactuals have no meaning. All that can happen does happen. This doesn't trouble him, however, because the multiverse is bigger than spacetime. ("To a first approximation, the multiverse is like a very large number of co-existing and

slightly interacting spacetimes.” (Deutsch 97, p. 275) So reality includes universes in which objective facts make our counterfactuals true.

A historian might make the judgment that ‘*if* Faraday had died in 1830, *then* technology would have been delayed for twenty years’....There is nothing arbitrary about which variants of our universe the counter-factual ‘*if* Faraday had died in 1830...’ refers to: it refers to *the variants which really occur* somewhere in the multiverse. That is what resolves the ambiguity. Appealing to imaginary universes does not work, because we can imagine any universes we like, in any proportions we like. But in the multiverse, universes are present in definite proportions, so it is meaningful to say that certain types of event are ‘very rare’ or ‘very common’ in the multiverse, and that some events follow others ‘in most cases’....Therefore the ‘*if...then...*’ statement can unambiguously be taken to mean ‘in most universes in which Faraday died in 1830, technological progress was delayed relative to our own.’ (Deutsch 97, p. 276)

The analysis being floated, integrating the multiverse with Nozick’s and McGinn’s accounts of knowledge, is not an argument for the truth of the multiverse hypothesis or a declaration of its truth. Rather, the possibility that it is true recommends a satisfying integration of McGinn’s and Nozick’s versions of reliabilism and what may be our best science about the cosmos. That the proposal is not idle speculation is indicated by its standing among string theorists and quantum cosmologists. For instance, Michael Clive

Price's *The Everett FAQ* reports that "political scientist" L David Raub polled 72 of the "leading cosmologists and other quantum field theorists" about the "Many-Worlds Interpretation." Raub gave the following response breakdown.

1. "Yes, I think MWI is true" 58%
2. "No, I don't accept MWI" 18%
3. "Maybe it's true but I'm not yet convinced" 13%
4. "I have no opinion one way or the other" 11%

Price writes

Amongst the "Yes, I think MWI is true" crowd listed are Stephen Hawking and Nobel Laureates Murray Gell-Mann and Richard Feynman. Gell-Mann and Hawking recorded reservations with the name "many-worlds", but not with the theory's content. Nobel Laureate Steven Weinberg is also mentioned as a many-worlder, although the suggestion is not when the poll was conducted, presumably before 1988 (when Feynman died). The only "No, I don't accept MWI" named is Penrose.

The findings of this poll are in accord with other polls, that many-worlds is most popular amongst scientists who may rather loosely be described as string theorists or quantum gravitists/cosmologists. It is less popular amongst the wider scientific community who mostly remain in ignorance of it.

This is enough perhaps to recommend a tracking analysis, at least in the spirit of Nozick's conception of philosophical explanation as opposed to coercive proof.

Starting from the possible-worlds construal that Nozick and McGinn both accept, these worlds are interpreted as the physically possible, parallel worlds of the multiverse.

Subjunctives now have categorical bases, as McGinn requests; they are grounded in an actual world, one that exists just as ours does, in which the antecedent is true. The subjunctive form of the tracking theory is vindicated as an interim analysis. And its subjunctives help define the discriminative capacities that figure in McGinn's version, which may now be called the third tier of the tracking theory of knowledge.

This tier's serious work is what McGinn envisages for it. Its key notion of discriminative capacity could unify propositional knowledge with knowing how, knowing who, and other forms of knowledge. Propositional knowledge emerges as fundamental (contrary to McGinn's expectation on this point) because of its role in the basic, "categorical" possible-worlds tier and in the second-level or "interim" tier given by Nozick's subjunctive statement of the tracking theory.

Not every objection levelled by McGinn against Nozick has been taken up here. In particular, Nozick's treatment of mathematical knowledge has not been defended. (It drops the variation condition because hypothesizing the falsehood of mathematical propositions, as in "If it were not the case that $2 + 2 = 4$...", may fail to make sense.) On the other hand, if McGinn's discriminatory-capacity version of reliabilism is correctly framed as the third tier of the tracking theory, its capacities may be called upon to explain how mathematical knowledge works, not by way of confounding the tracking theory but by contributing to an improved version of it.

The most troubling question that has been more begged than answered here is about the relationship between subjunctives or counterfactuals, on one hand, and capacities, capabilities, abilities, powers, and so on, on the other hand. The three-tier version of the tracking theory implies that subjunctives are logically prior to capacities, whereas McGinn urges the opposite, as in the following passage.

In the first place, it seems to me that this explication [of capacities by reference to counterfactuals gets the logical priority the wrong way around; for I would hold, quite generally, that an ascription of capacity is what grounds the associated counterfactuals—it is not that the capacity ascription is true in virtue of the truth of the counterfactuals. This claim is, I think, just a corollary of the general position about counterfactuals and categoricals that I allied myself with earlier: the counterfactuals about what someone would do in such-and-such circumstances are true *because (inter alia)* the person has a certain capacity—the person does not have the capacity because he satisfies the associated counterfactuals. (McGinn 1999, p. 16)

As McGinn acknowledges, he didn't establish this view. Nor has its denial been established here. However, once doubts about subjunctives have been assuaged by securing them in multiverse worlds, they give determinate content to capacity ascriptions that would otherwise be intolerably vague. To assert that Puck can be an ass might mean many things, specified by subjunctives about his rude behavior at parties, his transmogrifying by magic, his wearing a papiêr-maché donkey's head, and so forth. So the tracking theory's subjunctives should be accepted as more fundamental. Capacities are specified by subjunctives, and subjunctives are grounded in possible worlds.

The foregoing three-tier reliabilist account of *S knows that p* leaves hostages to scientific

fortune: the multiverse interpretation of possible-worlds semantics could turn out to be as imaginary as *A Midsummer Night's Dream*. But a certain vulnerability is becoming in a philosophical analysis. With reference to integration of knowledge with the multiverse, the vulnerability reveals a further implication of the analysis: What distinguishes true counterfactuals from mere fantasy is the grounding of the former in parallel worlds. If there is no physically possible world in which the liquor of the flower love-in-idleness, placed upon someone's brow, causes that person to madly dote on the next live creature that it sees; then that herb's powers are simply Shakespeare's fancy. If all counterfactuals are similarly bereft of grounding, we are not knowers; we do not track. What could or might be is what actually is, and no more. Knowledge is impossible.

The integration of knowledge and the multiverse has been presented as an illuminating structure, not as something proven. For all that has been *shown* here, it can be viewed as a denial that knowledge is possible as well as a statment of what knowledge requires. Whatever the wattage of the illumination, the structure is small. It does not canvas important objections to the tracking theory such as Christopher Peacocke's (1986), nor does it scout the prospects of integration with other forms of reliabilism, such as Alvin Goldman's development of it (1967, 1976). The hypothesis for a larger project would be that distinguishing different methods, along the lines of the reply to McGinn, would answer objections like Peacocke's; and that other forms of reliabilism would integrate readily with tracking and multiverse-possible worlds, especially those like Goldman's

that bring counterfactuals into the analysis rather than shunning them in McGinn's fashion.

REFERENCES

- Brown, Julian. 2001. *The Quest for the Quantum Computer*. New York: Simon and Schuster.
- Deutsch, David. 1997. *The Fabric of Reality*. New York: Penguin.
- Deutsch, David. 1985. "Quantum Theory, the Church-Turing principle and the universal quantum computer", in *Royal Society of London*, Volume A 400.
- Goldman, Alvin. 1976. "Discrimination and Perceptual Knowledge", in *Journal of Philosophy*, Volume 73.
- Goldman, Alvin. 1967. "A Causal Theory of Knowing", in *Journal of Philosophy*, Volume 64.
- Lewis, David. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
- McGinn, Colin. 1999. *Knowledge and Reality*. Oxford: Clarendon Press.
- Norris, Christopher. 2000. *Quantum Theory and the Flight from Realism*. London: Routledge.
- Norris, Christopher. 1999. "Should Philosophers Take Lessons from Quantum Theory? ", in *Inquiry*, Volume 42.
- Nozick, Robert. 1981. *Philosophical Investigations*. Cambridge, MA: Harvard University Press.
- Peacocke, Christopher. 1986. *Thoughts: An Essay on Content*. Oxford: Blackwell.
- Price, Michael Clive. 1995. "The Everett FAQ", <http://hedweb.com/everett/everett.htm>.
- Williams, Michael. 2002. "Nozick on Knowledge and Scepticism", in *Robert Nozick*. Cambridge,

MA: Cambridge University Press.

Copyright © 2004 Minerva

All rights are reserved, but fair and good faith use with full attribution may be made of this work for educational or scholarly purposes.

Wesley Cooper teaches Philosophy at the University of Alberta, Edmonton.

Email: amvu@shaw.ca