Frame Text:

Process Reliabilism for Justification

The first reliabilist approach to justification, and the one most widely discussed, is process reliabilism. This was originally formulated by Goldman in "What Is Justified Belief?" (1979). Goldman begins by proposing some constraints or desiderata for any account of justification. First, theories of justification should specify conditions for justified belief that do not invoke the justification concept itself, or any other epistemically normative concepts such as reasonability or rationality. The aim—or hope, at any rate—is to provide a "reductive" account of justification that doesn't rely, explicitly or implicitly, on any notions that entail justification or other members of the same family. This requirement has bite to it. For example, it might p¬reclude an analysis of justified belief in terms of "evidence", unless "evidence" can itself be characterized in non-epistemic terms. What kinds of terms or properties are appropriate, then, for constructing an account of justification? Permissible concepts or properties would include doxastic ones, such as belief, disbelief and suspension of judgment; and any other purely psychological concepts, such as ones that refer to perceptual experience or memory. Given the assumption that truth and falsity are non-epistemic notions, they would also be perfectly legitimate for use in analyzing justifiedness. Another admissible element in an account of justifiedness, it was proposed, is the causal relation.

Proceeding under these constraints, Goldman was led to the reliable process theory as follows. (The main theory is addressed to doxastic justifiedness—i.e., having of a justified belief—rather than propositional justifiedness—i.e., having justification for a proposition. It will be the sole topic of our discussion.) First, examples were used to show that whether or not a particular belief is justified depends on how that belief is caused, or causally sustained. Suppose that Sharon believes (justifiedly) a conjunction of propositions, Q and R, from which P logically follows. And suppose that, soon after forming this conjunction of beliefs, Sharon also forms a belief in P. Does it follow that Sharon's belief in P is justified? No. First, although Sharon believes Q and R, those propositions may play no (causal) role in her coming to believe P. She may form her belief purely by wishful thinking. She hopes that P is going to be true, and therefore (somehow) comes to believe it. Alternatively, suppose she uses some kind of "reasoning" that begins with Q and R. It is quite confused reasoning but serendipitously leads to P. In neither case is her resulting belief in P justified. This shows that a necessary condition

1

on a belief to be justified is that it be produced or generated in a suitable way. What kinds of belief–forming processes are suitable or proper, and what kinds are defective or unsuitable?

One feature that wishful thinking and confused reasoning have in common is unreliability. By contrast, which types of belief-forming processes confer justification? They include standard perceptual processes, remembering, good reasoning, and introspection. What do these processes share? Reliability: most of the beliefs they produce are true. (This formulation is slightly refined later.) Thus, the main proposal of "What Is Justified Belief?" was that a belief's justifiedness is fixed by the reliability or unreliability of the process or processes that cause it. Reliability might be understood in a frequency sense (pertaining to what occurs in the actual world) or a propensitysense (pertaining both to actual-world and possible-world outcomes). Justification is conferred on a belief by the truth-ratio (reliability) of the process that generates it. Just how high a truth-ratio a process must have to confer justification is left vague, just as the justification concept itself is vague. The truth-ratio need not be 1.0, but the threshold must surely be greater (presumably quite a bit greater) than .50.

A number of consequences were inferred from these main points, and refinements were added. One consequence was that process reliabilism is a "historical" theory. A reliable inferential process, for example, confers justification on an output belief only if the input beliefs (premises) are themselves justified. How could their justifiedness have arisen? Presumably, by having been caused by earlier applications of reliable processes. This chain must ultimately terminate in reliable processes that themselves have no doxastic inputs. Perceptual inputs are a good candidate for such processes. Thus, on this approach, justifiedness is often a matter of a history of personal cognitive processes. The historical feature of process reliabilism contrasts sharply with traditional foundationalism and coherentism, in which one's concurrent mental states are the only justification-determining factors.

These fundamental ideas were spelled out by Goldman in "What Is Justified Belief?" (1979/2012) in a series of principles: base-clause principles and recursive-clause principles. The initial one was (1):

- (1)If S's believing p at t results from a reliable cognitive belief-forming process (or set of processes), then S's belief in p at t is justified.

This principle may fit cases of perceptually caused beliefs and other beliefs that make no use of prior doxastic states (as inputs), but inferential beliefs seem to require a different principle.

When a belief results from inference, its justificational status depends not only on the properties of the inferential process but also on whether the premise beliefs of the inference are themselves justified. To accommodate this, a slightly more complex principle was introduced:

•   (2)If S's belief in p results from a belief-dependent process that is conditionally reliable, and if the beliefs (if any) on which this process operates in producing S's belief in p at t are themselves justified, then S's belief in p at t is justified.

By philosophical standards, these are not terribly complex principles; and, perhaps they invoke only a smallish set of core ideas. Thus, process reliabilism is a comparatively simple and straightforward theory. Such simplicity has usually been viewed as a virtue of the approach. (After all, theories in this territory are trying to capture the intuitive conception of justification of ordinary folk. How complex can their conception be? Ceteris paribus, then, simple theories are preferable to more complex ones.) Of course, matters are more complicated than the foregoing principles convey. They ignore cases in which the agent has "defeating" evidence for the proposition he or she comes to believe. "What Is Justified Belief?" therefore proposed a further principle to accommodate this additional detail. But we shall not explore this further complication (to appreciate its significance, however, see Beddor 2015).

The attractiveness of reliabilism can be illustrated by seeing how it handles a challenging type of example. How might it handle directly justified beliefs, for instance? Richard Feldman (2003) presents the following case. Two bird-watchers, a novice and an expert, are together in the woods when a pink-spotted flycatcher alights on a branch. Both form a belief that it's a pink-spotted flycatcher. The expert is immediately justified in believing this but the novice isn't; the latter just jumps to this conclusion out of excitement. Process reliabilism has adequate resources to handle this case (Goldman 2008). The crucial difference between expert and novice lies in the difference between their respective belief-forming processes. The expert presumably connects selected features of his current visual experience to things stored in memory about pink-spotted flycatchers, securing an appropriate match between features in the experience and features in the memory store. The novice does no such thing; he just guesses. Thus, the expert's method of identification is reliable whereas the novice's method is unreliable.

Because of the influence—though hardly uncontested influence—of this work on process reliabilism, as well as the reliabilist work surveyed in section 1, many commentators see epistemology as having undergone a major shift in recent decades. Michael Williams writes:

Since the nineteen sixties, Anglophone epistemology has undergone a paradigm shift: "the Reliabilist Revolution". (Williams forthcoming)

Williams himself seeks to resist this revolution, but does not dispute its occurrence. To pinpoint the core changes, it helps to distinguish two types of approaches to justification: "internalism" and "externalism". Internalism is usually identified as the dominant theme in epistemology's history since Descartes, continuing through most of the 20th century. Externalism is the new game in town, of which reliabilism is a salient example. What are the main features of internalism and externalism respectively?

There are two ways to fix what properties or states of affairs qualify as justifiers, or J-factors, according to internalism. On one option, a property or state of affairs F is a justifier for agent S(at t) only if F is directly accessible to S at t. On the second option, a property or state of affairs F is a justifier for S at t only if F is a mental state of S at t. The first view is called "accessibilism" and the second "mentalism" (Feldman and Conee 2001). What is direct accessibility? Roughly, it means knowability by some introspective or reflective method. Externalism is, generally speaking, the denial of internalism. For present purposes, it is a denial of both of the indicated forms of internalism. Given these definitions, it is pretty obvious that reliabilism must be a variety of externalism, because it holds that being caused by a reliable process is a (prima facie) justifier of a belief. But being reliably caused is neither a (pure) mental state nor something directly accessible in the intended sense. Being reliably caused is a matter of truth-conduciveness, and truth conduciveness is not introspectively or reflectively accessible. Similarly, reliabilism holds that processes used in the past may be justificationally relevant to a currently held belief (because of the historicity of justifiedness). But processes used in the past are not mental states concurrent with the target belief, and are not, in general (if at all), directly accessible to an agent now. Thus, since reliabilism doesn't require these properties to hold of J-factors, it is a form of externalism. Of course, reliabilists don't shy away from this consequence; they are generally happy with these features of their package. First, reliability theories avoid the stringent conditions of some internalist theories, conditions that are arguably too demanding to account for the extent of justified belief. Second, examination of cases

strongly supports process reliability as central to justification. Thus, although tweaking may be needed, reliabilists see externalism as a good path for epistemology to follow.

Although the details of the internalism/externalism debate are complex—and won't be pursued further here [see SEP entry on internalist vs. externalist conceptions of epistemic justification]—it should be clear that there is a major dispute here, so that a departure from the internalist perspective, as process reliabilism advocates, is indeed a substantial matter for epistemologists. Hence Williams's talk of a "revolution".[1]

3. Problems for Process Reliabilism

A number of problems for process reliabilism were identified in its own initial formulation or shortly thereafter. One type of problem is that its conditions seem too weak for justifiedness. Does it suffice for a belief's justifiedness that it be caused by a reliable process? Mustn't it also meet a meta-justification condition, for example, a "J→JJJ→JJ" condition, according to which if one's belief in p is justified, then one also justifiedly believes that one justifiedly believes p? Explicit use in a theory of the JJ principle itself, of course, would violate the constraints for a reductive account of justification. An account of justification (or at least a "base-clause" component of such an account) should not feature the very notion of justification itself. All right, but maybe one could add a requirement that the agent have a reliably-caused higher-order belief that his/her first-order belief is reliably caused. This proposal, unfortunately, is both too strong and too weak. It is too strong because agents do not constantly monitor their first-order beliefs for reliability and form higher-order beliefs about them. To require such continual monitoring as a condition of first-order justifiedness would be excessive. Too few beliefs would qualify as justified. Second, if one feels the need for higher-level requirements, why should they stop at the second level? Why not require a third-order reliably formed belief, and a fourth-order one, etc.? Here looms the threat of an infinite regress. Third, why should a critic who regards simple reliable causation as insufficient for justification be satisfied with any higher-order requirements? If simple reliable causation at the first level is insufficient, why should justification be guaranteed by reliability at any higher level? Some reliabilists will be inclined to strengthen the requirement for justification by adding a negative requirement, namely, that the agent not believe that her first-order belief is unreliably caused (or—what is arguably more in keeping with the spirit of reliabilism—that the agent not reliably believe that her first-order belief is so-caused).

A second problem for process reliabilism is the "new evil-demon problem" (Cohen 1984; Pollock 1984; Feldman 1985; Foley 1985). Imagine a world where an evil demon creates non-veridical perceptions of physical objects in everybody's minds. All of these perceptions are qualitatively identical to ours, but are false in the world in question. Hence, their perceptual belief-forming processes (as judged by the facts in that world) are unreliable; and their beliefs so caused are unjustified. But since their perceptual experiences—hence evidence—are qualitatively identical to ours, shouldn't those beliefs in the demon world be justified? Evidently, then, reliabilism must deliver the wrong verdict in this case.

One line of response to this problem is to argue that it doesn't follow from the low truth-ratio of processes in the demon world that the beliefs must be categorized as unjustified according to reliabilism, because reliabilism need not use the processes' truth-ratios in the world of the example as the standard of evaluation. That this is the standard was assumed in posing the objection; but it wasn't clearly so stated in the formulation of reliabilism. It is open to reliabilists to chart a different course, to choose a different standard of process reliability. Perhaps the appropriate domain or standard is the truth-ratio of the processes in the actual world. However, the plausibility or rationale for such an alternative standard is not obvious. We return to this issue in section 4 and again in sub-section 5.2.

A third objection to reliabilism, which also surfaced early on, argues that reliability isn't sufficient for justification. The principle example here is due to Laurence BonJour (1980). His strongest example describes a subject, Norman, who has a perfectly reliable clairvoyance faculty, but no evidence or reasons for or against the general possibility of a clairvoyant power or for or against his possessing one. One day Norman's clairvoyance faculty produces in him a belief that the President is in New York City, but with no accompanying perception-like experience, just the belief. Intuitively, says BonJour, he isn't justified in holding this belief; but reliabilism implies that he is. Similar examples were offered by Keith Lehrer (1990) and Alvin Plantinga (1993). We will re-visit these cases in section 4.

A fourth problem for reliabilism—perhaps the most discussed problem—is the generality problem. Originally formulated by Goldman in "What Is Justified Belief?", it has been pressed more systematically by Feldman (1985) and Conee and Feldman (1998). Any particular belief is the product of a token causal process in the subject's mind/brain, which occurs at a particular time and place. Such a process token can be "typed", however, in many broader or narrower ways. Each type will have its own associated level of reliability, commonly distinct from the

levels of reliability of other types it instantiates. Which repeatable type should be selected for purposes of assigning a reliability number to the process token? If no (unique) type can be selected, what establishes the justificational status of the resulting belief? Conee and Feldman (1998) lay down three requirements for a solution to the generality problem. First, the solution must be "principled" rather than ad hoc. Second, the type selected should have a reliability plausibly correlated with the justificational status of the ensuing belief. Third, the solution must remain true to the spirit of reliabilism. They argue, however, that prospects for finding such a solution are bleak.

A fifth problem, the problem of bootstrapping (or "easy knowledge"), is due to Jonathan Vogel (2000) and Stewart Cohen (2002). Roxanne is a driver who believes whatever her gas gauge "says" about the state of her fuel tank, although she has no antecedent reasons to believe it is reliable. Roxanne often looks at the gauge and arrives at beliefs like the following: "On this occasion the gauge reads 'F' and F", where the second conjunct implies that the tank is full. Since the perceptual process by which she arrives at the belief that the gauge reads 'F' is reliable, and so is the process by which she arrives at the belief that the tank is full (given that the gauge functions completely properly). According to reliabilism, therefore, her belief in the indicated conjunction should be justified. Now Roxanne deduces the proposition, "On this occasion, the gauge is reading accurately." And from (multiple examples of) this she induces "The gauge is reliable (in general)". Finally, with a little more deduction she concludes she is justified in believing that her gas tank is full. Since deduction and induction are reliable processes, Roxanne must also be justified in believing that her gas gauge is full. Suppose Roxanne does this repeatedly, without ever getting independent information about the gauge's reliability. Is she really justified in this? Definitely not, say Vogel and Cohen, because such bootstrapping amounts to epistemic circularity; it sanctions its own legitimacy (no matter what). So reliabilism gets this wrong.

A final problem (for present purposes) is the so-called "value problem". Plato claimed that knowledge is more valuable than true belief, and many authors concur with his suggestion. This raises the puzzle of why this should be so. What extra value does knowledge have as compared with true belief? Focusing on process reliabilism, the question is whether reliabilism can explain this value difference. (Although our present topic is justification, not knowledge, this organizational matter will be ignored.) Reliabilism's answer, it would seem, is that causation by a reliable process confers extra value on a belief so as to make it justified and/or knowledge. This suggestion is criticized by several philosophers: Jones (1997), Swinburne (1999),

Zagzebski (1996, 2003), Riggs (2002), and Kvanvig (2003). Zagzebski's example brings the point home. Consider a cup of espresso, she says, that is produced by a reliable espresso machine.

[T]he reliability of the source [the expresso machine] does not … give the product an additional boost of value. If the espresso tastes good, it makes no difference if it comes from an unreliable machine. (2003: 13)

Similarly, the epistemic value of a belief cannot be raised by the reliability of the source.