



# Stanford Encyclopedia of Philosophy

## Reliabilist Epistemology

*First published Mon Apr 21, 2008; substantive revision Wed Dec 2, 2015*

Reliabilism is an approach to epistemology that emphasizes the truth-conduciveness of a belief-forming process, method, or other epistemologically relevant factors. The reliability theme appears in theories of knowledge, of justification, and of evidence. “Reliabilism” is sometimes used broadly to refer to any theory that emphasizes truth-getting or truth indicating properties. More commonly it is used narrowly to refer to process reliabilism about justification. This entry discusses reliabilism in both broad and narrow senses, but concentrates on the theory of justification.

- [1. Reliability Theories of Knowledge](#)
- [2. Process Reliabilism for Justification](#)
- [3. Problems for Process Reliabilism](#)
- [4. Replies, Refinements and Modifications](#)
- [5. New Developments and Extensions](#)
  - [5.1 Normality Reliabilism](#)
  - [5.2 Beyond “Individualistic” Reliabilism](#)
  - [5.3 Reliabilism and Degrees of Belief](#)
- [6. Cousins and Spinoffs of Process Reliabilism](#)
  - [6.1 Virtue Reliabilism](#)
  - [6.2 Syntheses of Reliabilism and Evidentialism](#)
- [Bibliography](#)
- [Academic Tools](#)
- [Other Internet Resources](#)
- [Related Entries](#)

---

## 1. Reliability Theories of Knowledge

It is generally agreed that a person *S* knows a proposition *P* only if *S* believes *P* and *P* is true. Since all theories accept this knowledge-truth connection, reliabilism as a distinctive approach to knowledge is restricted to theories that involve truth-promoting factors above and beyond the truth of the target proposition. What this additional truth-link consists in, however, varies widely.

Perhaps the first formulation of a reliability account of knowing appeared in a brief discussion by F.P. Ramsey (1931), who said that a belief is knowledge if it is true, certain and obtained by a reliable process. This attracted no attention at the time and apparently did not influence reliability theories of the 1960s, 70s, or 80s. Another early reliability-type theory was Peter Unger’s (1968) proposal that *S* knows that *P* just in case it is not at all accidental that *S* is right about its being the case that *P*. *S*’s being right about *P* amounts to *S*’s *believing* truly that *P*. Its not being accidental that *S* is right about *P* amounts to there being something in *S*’s situation that makes it highly probable that *S* would be right. David Armstrong (1973) offered an analysis of non-inferential knowledge that explicitly used the term “reliable”. He drew an analogy between a thermometer that reliably indicates the temperature and a belief that reliably indicates the truth. According to this account, a non-inferential belief qualifies as knowledge if the belief has properties that are nomically sufficient for its truth, i.e., guarantees its truth via laws of nature. This can be considered a reliable-indicator theory of knowing. Alvin Goldman offered his first formulation of a reliable process theory of knowing—as a refinement of the causal theory of knowing—in a short paper on innate knowledge (Goldman 1975).

In the 1970s and 1980s several subjunctive or counterfactual theories of knowing were offered with reliabilist contours. The first was Fred Dretske's "Conclusive Reasons" (1971), which proposed that *S*'s belief that *P* qualifies as knowledge just in case *S* believes *P* because of reasons he possesses that would not obtain unless *P* were true. In other words, *S*'s reasons—the way an object appears to *S*, for example—are a reliable indicator of the truth of *P*. This idea was elaborated in Dretske's *Knowledge and the Flow of Information* (1981), which linked knowing to getting information from a source through a reliable channel. Meanwhile, Goldman also proposed a kind of counterfactual reliability theory in "Discrimination and Perceptual Knowledge" (1976). This theory developed the idea of knowledge excluding "relevant alternatives", an idea already adumbrated in Dretske's "Epistemic Operators" (1970). In Goldman's treatment, a person perceptually knows that *P* just in case (roughly) she arrives at a belief in *P* based on a perceptual experience that enables her to discriminate the truth of *P* from all relevant alternatives. On this approach, *S*'s knowing that *P* is compatible with there being "radical" (hence irrelevant) situations—for example, brain-in-a-vat situations—in which *P* would be false although *S* has the same experience and belief. Gail Stine (1976) explored this approach with respect to knowledge, skepticism, and deductive closure (i.e., the principle that one knows all that is implied—or all that one knows to be implied—by what one knows).

Robert Nozick (1981) proposed a theory with similar contours, a theory he called a "tracking" theory. In addition to truth and belief, Nozick's conditions for knowledge were: (1) if *P* were not true then *S* would not believe that *P*, and (2) if *P* were true, *S* would believe that *P*. The first of the two tracking conditions was subsequently called the "sensitivity" requirement. A number of counterexamples have been produced to this condition (see especially DeRose 1995). A similar tracking condition that has gained attention recently is a "safety" condition. Safety can be explained in slightly different formulations (see Ernest Sosa 1996, 2000; Timothy Williamson 2000; Duncan Pritchard 2005), including "if *S* believes that *P*, then *P* would not easily have been false", or "in all of the nearest worlds where *S* believes that *P*, *P* is true". Williamson classifies the safety approach as a species of reliability theory (2000: 123–124).

Reliability theories are partly motivated by the threat of skepticism. It is natural to think that if you know that *P* then in *some sense* you "can't be wrong" about *P*. But what is the relevant sense of "can't"? Does it mean that your evidence must *logically* preclude the possibility of error? If so, very few propositions would be known. Reliability theories, in their various ways, propose weaker but still substantial senses of "can't be wrong". For example, the relevant-alternatives theory allows that one can know that *P* even if there are logically possible situations in which one's evidence is the same but *P* is false. But it insists that there be no *relevant* possible situations in which one's evidence is the same but *P* is false. Such an account is not so seriously threatened by skepticism.

Reliability theories of knowledge continue to appeal to epistemologists, and permutations abound. The reliability theories presented above focus on *modal* reliability, on getting truth and avoiding error in possible worlds with specified relations to the actual one. They also focus on *local* reliability, that is, truth-acquisition in scenarios linked to the specific scenario in question as opposed to truth-getting by a process or method over a wide range of cases. Other reliabilisms focus on *global* reliability: the reliability of the type of process or method used across all or many of its applications. Goldman's *Epistemology and Cognition* (1986) combines both local and global reliability in its account of knowledge.

## 2. 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 preclude 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 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 *propensity* sense (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 20<sup>th</sup> 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 justificationaly 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](#)]<sup>[1]</sup>—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".<sup>[1]</sup>

### 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 " $J \rightarrow JJ$ " 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 espresso 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.

## 4. Replies, Refinements and Modifications

Reliabilists have offered a number of responses to these various problems and objections. Having already considered a response to the first problem (in [section 3](#)), we turn next to the second and third: the new evil-demon problem and the clairvoyance problem.

One response to the clairvoyance problem is to distinguish *prima facie* from *ultima facie* justification. Process reliabilism is typically couched as a theory of *prima facie* justification: if *S*'s belief that *p* is the result of a reliable process, but it is maintained in the face of (sufficiently strong) countervailing considerations, *S*'s belief will not count as *ultima facie* justified (that is, it will not be justified *full stop*). Thus one possibility is that Norman's belief that the president is in New York is not *ultima facie* justified because it is defeated. A response along these lines was briefly suggested by Goldman (1986: 112).

Of course, a response along these lines needs to be supplemented with a theory of defeat. Goldman's original 1979 article proposed that *S*'s belief is defeated as long as there are reliable (or conditionally reliable) belief-forming processes available to *S* such that, if *S* had used those processes in addition to the process actually used, *S* wouldn't have held the belief in question. However, it is not entirely clear how this "Alternative Reliable Process" account would help in the case of Norman; what's more, some have deemed this theory of defeat problematic on other grounds (see Beddor 2015 for objections). But even if the Alternative Reliable Process account is deemed wanting, presumably some superior account of defeat is possible. (Indeed, it seems that any adequate account of justification—not just those of a reliabilist bent—owes us some story about defeat.) And it may be that the right account of defeat—whatever it is—will help process reliabilism capture the intuition that Norman's belief is unjustified.

A different response to the clairvoyance objection is to concede that being the result of a reliable process isn't sufficient for even *prima facie* justification; some further condition must be met. Jack Lyons (2009, 2011) develops a novel reliabilist response to the clairvoyance challenge along precisely these lines. According to Lyons, in order for a non-inferential belief to be justified, it must be the result of a "primal system".

Drawing on research in cognitive science, Lyons proposes that a primal system is any cognitive system that meets two conditions: (i) it is "inferentially opaque"—that is, its outputs are not the result of an introspectively accessible train of reasoning, (ii) it develops as a result of a combination of learning and innate constraints (2009: 144). For Lyons, our perceptual systems are paradigmatic examples of such primal systems.

How does this help with the clairvoyance objection? According to Lyons, Bonjour's presentation of the Norman example invites the assumption that Norman's clairvoyance was the result of some recent development—e.g., "a recent encounter with radioactive waste" or a "neurosurgical prank"—not the result of some combination of learning and innate constraints (2009: 118–119). Given this assumption, Norman's clairvoyance-based belief is not the result of a primal system, hence it is not *prima facie* justified. Lyons goes on to argue that if we consider variants of the Norman case where the agent's belief is the result of a primal system, our intuitions shift. Lyons asks us to consider the case of Nyrmoon, a member of an alien species, for whom clairvoyance is a normal cognitive capacity. (To make this more plausible, Lyons asks us to imagine that Nyrmoon—and his conspecifics—can detect "highly attenuated energy signals from distal events".) However, Nyrmoon is so unreflective that he has no beliefs about the reliability of his clairvoyance. Lyons contends that, in contrast with Norman, Nyrmoon's clairvoyance-based beliefs are justified. He takes this to support the claim that being the result of a primal system that is also reliable is sufficient for the justifiedness of noninferential beliefs.

Thus far we have looked at responses that focus on the clairvoyance objection. Another strategy is to try to solve both the new evil demon problem and the clairvoyance problem at one fell swoop by opting for a variant of process reliabilism. This variant, originally called "two-stage reliabilism" (Goldman 1992), has been endorsed more recently under a more appealing label: "approved-list reliabilism" (Fricker forthcoming). The approach is inspired by the following conjecture about how attributors make justification attributions. In a preliminary stage, opinions are formed about the reliability of assorted belief-forming processes, using observation and/or inference to draw conclusions about the track-records of these processes in the actual world. They thereby construct mental lists of reliable and unreliable processes: lists of approved and disapproved processes (respectively). In the second stage of the operation, they deploy these lists to make judgments about particular beliefs (actual or hypothetical). If somebody's belief was caused by a process that is on their approved list—or resembles one on their approved

list—they consider it justified. If it is caused by a process on their disapproved list, it is classed as unjustified.

How would approved-list reliabilism explain intuitive judgments in the clairvoyance case? Presumably, an ordinary attributor would not have a clairvoyance process on either of her lists. But she might well have processes like extra-sensory-perception or telekinesis on her list, especially her disapproved list. The process or faculty Norman uses to arrive at his belief about the President sounds very similar to one of those obscure and suspect powers. Hence, Norman's belief is intuitively classified as unjustified. This is despite the fact that—as potential attributors are told—Norman's clairvoyance process is thoroughly reliable. This is how approved-list reliabilism explains our judgments in the clairvoyance case. What about the new evil-demon case? Again, it is assumed as background that a potential attributor constructs lists of belief-forming processes, one for approved process types and one for disapproved types. Perceptual processes (of various sorts) would be on the approved list. Since the people in the evil-demon case use perceptual processes that would be on the approved list, an attributor would consider their resulting beliefs justified—even though s/he is told that the perceptual processes in the evil-demon world are unreliable. The whole idea of approved-list reliabilism is that the two stages—list construction and list application—are quite distinct. This enables the theory to predict responses to justification questions that comport with our actual inclinations or intuitions.

Approved list reliabilism is a theory of the factors that influence our *attributions* of epistemic justification. It is thus naturally construed as an *attributor theory*—a theory of the conditions under which a justification attribution (that is, a sentence of the form, “*S* is justified in believing *p*”) is judged to be true or false. In this regard, it parallels attributor theories of knowledge (e.g., DeRose 1992, 2009). There are different ways of developing an approved-list attributor theory more precisely. For instance, a contextualist implementation might hold that a justification attribution is true if and only if the subject's belief-forming process belongs to the speaker's approved list. Alternatively, one could adopt an assessor-relativist implementation, according to which the truth-conditions of justification attributions are relativized to contexts of assessment (cf. MacFarlane 2005). An assessor-relativist version of the approved list might hold that a justification attribution is true at a context of assessment if and only if the subject's belief-forming process belongs to the assessor's approved list.

A similar strategy for solving the New Evil Demon problem is to distinguish two different ways a process can be reliable. Suppose we inhabit the actual world (@), and we're evaluating a subject *S* who inhabits some other world  $w_s$ . Then there are two different things we could mean when we say that *S*'s belief-forming process is reliable: we could mean that it's reliable relative to @, or we could mean that it's reliable relative to  $w_s$  (Sosa 1993, 2001). Comesaña (2002) uses this distinction to provide a solution to the new evil demon problem cast in the framework of two-dimensional semantics (Stalnaker 1999). On Comesaña's proposal, the sentence, “*S* is justified in believing *p*” has two readings: (i) that *S*'s belief-forming process is reliable relative to @, (ii) that *S*'s belief-forming process is reliable relative to  $w_s$ . If  $w_s$  is a demon world, then reading (i) will be true, but reading (ii) is false. While Comesaña's use of two-dimensional semantics has drawn recent criticism (Ball and Blome-Tillman 2013), the basic strategy of solving the New Evil Demon problem by appealing to two types of reliability remains popular. (A different promising approach to the New Evil Demon Problem, the “normality approach” is presented in [section 5.2](#) below.)

The fourth problem posed in [section 3](#) is the generality problem. Epistemologists who worry about process reliabilism because of the generality problem usually assume that a “solution” to this problem will consist in a formula for identifying a unique process type given any specified case (assuming the case is specified in reasonable detail). That is, pretty clearly, what Conee and Feldman mean to require. Some epistemologists have tried to provide such a formula, and some of them sound on the right track (e.g., Alston 1995; Beebe 2004). James Beebe, for example, says that a pertinent type will, first of all, be an information-processing procedure or algorithm. Of course, there will often be indefinitely many types of this kind, of varying reliability. To pick out the appropriate type, Beebe offers the following instructions. Let *A* be the broadest such type. Choose a partition that is the broadest objectively homogeneous subclass of *A* within which the token process falls, where a class is objectively homogeneous if no statistically



relevant partition of  $S$  can be effected. This is an interesting idea, but there remains the lingering worry that there may always be a set of different conditions that meet Beebe's standards, not a unique one.

A different kind of approach is less optimistic about finding a formula to pinpoint a unique type for each process token. This approach is less sanguine about what the philosopher can do here. However, this need not put any special pressure on process reliabilism. It may be a problem that faces any theory of doxastic justifiedness, at least any theory that highlights the causal provenance of a target belief. (And this is probably all viable theories.) This is the tack taken by Comesaña (2006). He argues that the generality problem is not a *special* problem for reliabilism, but one shared by all epistemologies of justification, including Feldman and Conee's own evidentialist theory. As Comesaña notes, every adequate epistemology needs an account of the basing relation, and any attempt to explain the basing relation ultimately runs into the generality problem, or something very similar to it. Michael Bishop (2010) argues for the same conclusion on different grounds. According to Bishop, the generality problem will arise for any theory that allows for the possibility of reflective justification—that is, having a belief  $B$  that is justified on the basis of one's knowledge that one formed  $B$  via a reliable form of reasoning.

But if no formula is proposed for selecting the uniquely correct process type, isn't it a mystery how attributors converge on the same classifications (justified vs. unjustified) a large percentage of the time? Yet there does seem to be such convergence. (Think about all of the easy cases, not the hard ones.) Can this be explained? If so, will the explanation sit comfortably with process reliabilism? Erik Olsson (forthcoming) calls our attention to a well-supported psychological theory about conceptualization called *basic-level* theory. It is mainly due to the work of Eleanor Rosch in the 1970s (Rosch et al. 1976). Rosch and her collaborators studied the deployment of taxonomically related concepts like "animal", "dog", and "Labrador". In such a taxonomy, one term is a superordinate concept ("animal"), another is an intermediate-level concept ("dog"), and a third is a subordinate concept ("Labrador"). It turns out that intermediate-level concepts have a privileged status, and are therefore called "basic-level" concepts. Basic-level concepts are overwhelmingly preferred in free naming, are the first concepts acquired by children, and occur more frequently in text. It was also demonstrated that people tend to converge on reports in using basic-level concepts. In one study psychologists found that, out of 540 responses, 530 to 533 converged on the same word for naming a physical object (Rosch et al. 1976). Olsson suggests that basic-level theory calls into question Feldman and Conee's contention that in the absence of contextual cues, there is no single intuitive process-type to which a given process token corresponds. In a related vein, Jönsson (2013) showed subjects clips in which characters arrived at various conclusions, and then asked the subjects to specify how the characters arrived at their beliefs. Subjects converged on the choice of verbs describing the belief-formation processes, even without linguistic cues to guide the process-typing task. A follow-up experiment (Jönsson 2013) found a correlation between subjects' estimates of the reliability of the characters' belief-forming processes and subjects' judgments about whether the characters were justified in holding the beliefs in question. Thus there is some evidence that folk psychological propensities lead us to converge on belief-typing tasks, and that our reliability assessments track our justification judgments.

Consider now the fifth and sixth problems of [section 3](#): the bootstrapping and extra-value problems. Brief responses will be given to each. As Cohen (2002) points out, reliabilism is not unique in facing the bootstrapping problem. Indeed, it is faced by all theories that exhibit what he calls "basic knowledge" structure. Moreover, as van Cleve (2003) effectively demonstrates, theories that do not allow for basic knowledge lead to (wide-ranging) skepticism.

If we wish to allow for basic knowledge, can we still give a principled explanation of why some forms of bootstrapping seem illegitimate (for example, the case of Roxanne)? This is an area of active research. One suggestion is that illegitimate forms of bootstrapping involve *No Lose Investigations*. Roughly, a No Lose Investigation into a hypothesis  $h$  is an investigation that could never, in principle, count against  $h$ . (For suggestions along these lines, see Kornblith 2009; Titelbaum 2010; Douven and Kelp 2013.) Another suggestion is that illegitimate forms of bootstrapping all involve *epistemic feedback* (Weisberg 2010). Suppose an agent believes premises  $p_1 \dots p_n$ , from which she infers lemmas  $l_1 \dots l_n$ , from which she in turn infers a conclusion  $c$ . Epistemic feedback is present when the probability of  $c$  conditional on  $l_1 \dots l_n$

is greater than the probability of  $c$  conditional on  $p_1 \dots p_n$ . Roxanne's case can be understood in these terms. She first believes various premises about the gas gauge readings (e.g., *The gas gauge read full at time  $t_1$ ; the gauge reads half-empty at  $t_2$* ). She then infers various lemmas about the state of the gas tank (e.g., *The tank was full at  $t_1$ ; the tank was half-empty at  $t_2$* ). Finally, by conjoining these premises with these lemmas, she comes to believe the conclusion: *The gas tank is reliable*. The probability of this conclusion conditional on just the lemmas (that is, the beliefs about the state of the gas tank) is higher than the probability of the conclusion conditional on the premises (that is, the gas gauge readings). Perhaps by imposing a ban on either No Lose Investigations or epistemic feedback (or both), we can account for the intuition about Roxanne, while still allowing for basic knowledge. (For an overview of various responses to the bootstrapping problem, see Weisberg 2012.)

The value problem is also one that may be questioned from the start. Was Plato right to claim that knowledge has more value than true belief? This is debatable. However, let us presume this value claim to be correct. A sophisticated account of value should be able to account for one true belief having more value than another true belief in virtue of its external relations to other events, specifically, the kind of process that generated it. A strong case is made for this—together with compelling examples—by Wlodek Rabinowicz and Toni Roennow-Rasmussen (1999; also see Rabinowicz and Roennow-Rasmussen 2003). What is of final value may nonetheless have this value for its own sake in part because of its relational properties (Kagan 1992).

One class of such things, discussed by Rabinowicz and Roennow-Rasmussen, involve cases in which a thing is valued for its own sake in virtue of its special relationship to another object, event, or person.

An original, say, an original work of art, may be valued for its own sake precisely because it has the relational property of being an original rather than a copy. Its final value supervenes, in part, on its special causal relation to the artist. Princess Diana's dress may be another case in point. The dress is valuable just because it has belonged to Diana. This is what we value it for.... [T]here are innumerable examples of a similar kind: Napoleon's hat, a gun that was used at Verdun, etc. In all these cases, a thing acquires a non-instrumental value in virtue of its causal relation to some person, object or event that stands out in some way (Rabinowicz and Roennow-Rasmussen 1999: 41).

Similarly, states of true belief that are instances of knowledge might have higher final value because they have the relational property of being caused by a reliable belief-forming process.

## 5. New Developments and Extensions

In this section we turn to new developments in reliabilist epistemology, novel proposals that appeared after the previous (2008) version of this entry was posted.

### 5.1 Normality Reliabilism

As we have already seen, a crucial question that faces reliability theories concerns the domain in which a process is assessed for reliability. Several of the most-discussed problem cases for reliabilism placed this question at center stage. In the new evil-demon case, for example, is it the world of the demon at which the reliability of perception is to be assessed, or is it the actual world (or the world of the assessor)? Recently a new type of answer has been floated by at least two writers. The new answer introduces the notion of *normal conditions* for the use of a given process or method, and suggests that reliability should be measured, or assessed, only in terms of how well it works in normal conditions (for the selected process or method).

For example, Jarrett Leplin (2007, 2009) rejects the common view of reliable processes/methods as those that produce a high ratio of truths to falsehoods. In its place, Leplin advances a conception of reliability according to which a process/method is reliable if it would never produce or sustain false beliefs under normal conditions. As Leplin notes, this is a “subjunctive version of reliabilism”, akin to Nozick's

treatment of knowledge (Leplin 2007: 33; 2009: 34–35).

How should we understand “normal conditions”? Leplin suggests:

[C]onditions normal for a method are conditions typical or characteristic of occasions and environments in which the method is usable or applicable, whether or not it is in fact then or there used or applied. My thermometer is not (for the most part) usable under extreme conditions. (Ordinary) perception fails in the dark. It might happen that on all actual occasions... of a method's use, the conditions are *atypical* of occasions on which it is usable. Then, despite the method's reliability, the preponderance of the beliefs it yields could be false. (2009: 37–38)

According to this account, perceptual beliefs of people in the demon world, for example, can now be judged to be justified because perceptual belief formation is a reliable process. It doesn't matter that perception isn't reliable in *their* environment; it suffices for the process to confer justification that it be reliable in *normal* worlds.

Christensen (2007) raises the question of whether Leplin's conception of reliability is too strong. He suggests that there are reliable processes (consulting the *Encyclopedia*, facial recognition) that sometimes yield false beliefs even when operating under normal conditions. One reason that it's difficult to judge the force of this objection is that the notion of “normal conditions”, even given Leplin's clarifications, remains vague (which is not obviously a problem with the analysis; cf. Leplin 2009). At the very least, Leplin's “normality” approach offers a promising variant of traditional reliabilism, though Leplin's full story (not recounted here) is quite complicated and not easily assessed.

An even more complex normality approach is presented by Peter Graham (2012). Graham draws on an etiological account of function due to Larry Wright (1973) and Ruth Millikan (1984), among others, to advance a theory of epistemic entitlement (= justification) in terms of proper functioning. A compact statement of his theory is this:

Epistemic entitlement, I argue, consists in correct or proper performance—or normal functioning—for belief-forming processes that have reliably forming true beliefs as an etiological function (2012: 449–450).

In a slightly different formulation,

[A] belief enjoys a kind or source of *prima facie* entitlement if and only if it is based on a normally functioning process that has reliably forming true beliefs as an etiological function. (2012: 472)

Given this basic theoretical framework, Graham draws a very similar conclusion for the brain-in-the-vat case as that drawn by Leplin for the demon-world case:

What the brain-in-a-vat case suggests is that entitlement persists even when not in normal conditions, as long as the process is functioning normally. (2012: 471)

Like Leplin, then, Graham uses a normality approach to address some familiar counterexamples to process reliabilism. Whether one or more of these normality approaches fully achieves the end at which it aims, which remains to be seen, they certainly introduce a fresh and well-motivated angle that holds promise for dealing effectively with familiar counterexamples.

## 5.2 Beyond “Individualistic” Reliabilism

Most versions of reliabilism are “individualistic” in at least two senses. First, they focus on the justificatory status of a belief of an individual agent. Second, they make the justificatory status of such a belief depend entirely on the reliability of processes that take place within her or his head. Recently, a number of authors

have explored versions of reliabilism that revise or abandon these individualistic assumptions.

Sanford Goldberg (2010) advances a distinctive view of testimonial belief that abandons the second individualistic assumption, which he calls, “Process Individualism”. On Goldberg’s view, a proper assessment of the justificatory status (hereafter, J-status) of a testimonial belief requires

an assessment of the reliability of cognitive processes that take place *in the mind/brain(s) of the subject’s informant(s)*. (Goldberg 2010: 2, emphasis original)

One of the ways that Goldberg motivates this “extendedness thesis” is via the following sort of case (chap. 4). An informant (*A*) forms a perceptual belief that *p*, which she conveys via testimony to an audience (*B*), who then comes to believe *p*. However, *A*’s perceptual belief that *p* was formed in a way that falls just shy of the threshold for justification, perhaps because it is based on a momentary glance, or formed in poor lighting conditions. Goldberg contends that *B*’s testimony-based belief that *p* does not amount to knowledge. What’s more, the reason why it does not amount to knowledge is that it is not justified. After all, the belief could well be true, and free from Gettierization. (For an overview of the Gettier Problem, see the entry on the Analysis of Knowledge.) But if *B*’s belief that *p* is not justified, this justificatory failing is not due to any unreliability in *B*’s mental processing of *A*’s testimony; it’s rather due to the fact that *A*’s original belief that *p* was insufficiently justified. Goldberg thus concludes that the J-status of testimonial beliefs are importantly affected by the reliability of the testifier’s cognitive processes.

A different new wrinkle in reliabilist epistemology has recently emerged in social epistemology. It considers a non-traditional kind of agent and asks what it takes for such an agent to have justified beliefs. Could process reliabilism be applicable here too? The kind of agent in question is a collective agent. Collective, or group, agents are “plural subjects”. They are entities treated as subjects in the sense that they are assumed to have propositional attitudes like desires, intentions, goals, beliefs, judgments, etc. We routinely speak this way in everyday discourse, and even in formal or legal contexts, in which corporate bodies are assumed to have the same kinds of attitudes that individuals have. We ask whether the CIA or the FBI “knew” that certain Al Qaeda perpetrators of the 9/11 attacks had been taking flight lessons in the United States. In asking this question, one assumes that the CIA and the FBI are the bearers of beliefs or information; in other words, they have a “psychology” of some sort (not to say *conscious* psychological states). Working in this framework, a number of philosophers pose the following question: under what circumstances does a group have a certain belief? It is widely assumed that the various belief-states of the group’s *members* are important determiners of whether the group as a whole has a belief with the same content. For example, Margaret Gilbert (1989) offers what she calls the “joint acceptance account” of group belief:

A group *G* believes that *p* if and only if the members of *G* jointly accept that *p*. The members of *G* jointly accept that *p* if and only if it is common knowledge in *G* that the members of *G* individually have intentionally and openly ... expressed their willingness jointly to accept that *p* with the other members of *G*. (1989: 306–307)

This account of group belief is not widely accepted, but this is incidental for present purposes. It is an early proposal for formulating conditions of group belief. The main issue discussed here, however, is not conditions for group *belief*, but conditions for the *justificational status* of group beliefs. This kind of question has been little discussed within epistemology. Only with the emergence of social epistemology—and, more specifically, with a recognition of the *collective* branch of social epistemology—has it come clearly into focus.

The idea that process reliabilism might be a viable approach to collective justifiedness was first broached by Frederick Schmitt (1994), but with considerable tentativeness. A more developed reliabilist approach is presented in Goldman 2014.

Goldman’s (2014) proposal makes use of List and Pettit’s 2011 framework for describing group beliefs, which we will briefly review. It is widely thought that a group’s beliefs are determined—in some way—by the individual beliefs of the group members (List and Pettit 2011: 64). How exactly does this

determination work? Different candidate answers to this question can be represented by different belief (or judgment) aggregation functions (BAFs), which take as inputs beliefs of individual agents, and produce as outputs group beliefs. A burgeoning literature is devoted to discussing plausible requirements on BAFs. For instance, one might require: (1) A BAF admits as input any possible profile of individual beliefs (doxastic attitudes) toward the propositions on an agenda, and the individual attitudes are consistent and complete (“universal domain”). (2) The BAF produces as output consistent and complete group attitudes towards the propositions on the agenda (“collective rationality”). (3) All individuals’ attitudes are given equal weight in determining the group attitudes (“anonymity”). And so on. Although each of these (plus additional) conditions seem initially plausible, a number of impossibility theorems have been proved that various combinations of such desiderata cannot be jointly satisfied. This raises many intriguing problems, which are not explored here.

Nothing has been said thus far about justification. What we might seek is something analogous to a BAF, namely, a *justification* aggregation function (JAF). Assuming that group  $G$  forms a belief in  $p$  in conformity with some BAF and its members’ beliefs *vis-à-vis*  $p$ , what must be true of the J-statuses of those members’ beliefs (and other attitudes) with respect to  $p$  for it to be true that  $G$ ’s belief is justified? If *none* of its members’ beliefs are justified, then presumably  $G$ ’s belief won’t be justified either. A larger question is how—or whether—the justifiedness of members’ beliefs is automatically transmitted toward the group’s justifiedness.

Goldman argues for treating this transmission on the model of inference *within* an individual agent. In the individual case, premise beliefs generate beliefs in a conclusion, and the J-status of the conclusion belief depends on the J-statuses of the premise beliefs. Special epistemic relations hold between states and epistemic statuses within a single individual. The fact that *another* person is justified in believing a proposition  $Q$  doesn’t give *me* justification for believing something implied by  $Q$ . What about relations between a group and its members? The suggestion is made that this is (or can be) an intimate relation. Just as justifiedness can be transmitted intra-individually, it can be transmitted within a group and its membership. Process reliabilism can then be brought in to capitalize on this parallel. According to process reliabilism for individuals, inferential justification depends on two factors: (a) the justifiedness of the premise beliefs and (b) the *conditional reliability* of the inferential process used. In the collective belief case group justification depends on two analogous factors: (a) the justifiedness of the members’ beliefs, and (b) the conditional reliability of the JAF: the function that maps member justifiedness into group justifiedness.

The analogy between the individual and the aggregative cases is not perfect, however. In cases of an individual’s inference, it is plausible to require *all* premise beliefs to be justified in order for the conclusion to be justified. This is too strong a requirement for the collective case, however. Surely a group can attain justified belief even if, say, only 95% of the members believe it justifiedly. Goldman handles group justifiedness, in the end, by talking about *degrees* of justifiedness for a group, arising from assorted distributions of member justifiedness. In particular, the following principle is proposed:

(GJ) If a group belief in  $P$  is aggregated based on a profile of member attitudes toward  $P$ , then *ceteris paribus* the greater the proportion of members who *justifiedly believe*  $P$  and the smaller the proportion of members who *justifiedly reject*  $P$ , the greater the group’s *grade of justifiedness* in believing  $P$  (Goldman 2014: 28).

With this as background, the idea is introduced that group justifiedness can arise through a diachronic process, in which members’ severally acquire justified or unjustified beliefs in the target proposition in accord with process reliabilist principles for individuals. The J-statuses of these individual beliefs then determine—via process reliabilist principles—the J-statuses of the group beliefs.

### 5.3 Reliabilism and Degrees of Belief

Historically, reliabilism has been offered as an account of the justificatory status of full or outright belief. However, it’s widely thought that beliefs come in degrees: a person might believe that it’s sunny and also



believe that it's Monday, but have a higher degree of belief in the former than the latter. This raises the question: can reliabilism be extended to provide an account of the justificatory status of degrees of belief?

Formal epistemologists and decision theorists have long been interested in different “scoring rules”—functions that measure the accuracy or inaccuracy of degrees of belief (hereafter, credences). For example, one widely discussed scoring rule is the *Brier score* (Brier 1950). Let  $C(p)$  be an agent's credence in  $p$ ; let  $T(p)$  be  $p$ 's indicator function, which equals 1 if  $p$  is true, and 0 if  $p$  is false.  $C(p)$ 's Brier score is calculated by the formula:

$$(C(p) - T(p))^2$$

Thus a credence of 1 in a true proposition will get a Brier score of 0—the best score possible. A credence of 1 in a false proposition will get a Brier score of 1—the worst score possible. An intermediate credence of .6 will get a Brier score of .16 if the proposition is true, and .36 if it is false.

Given a particular scoring rule  $R$ , we can develop a measure of process reliability (Dunn 2015; Tang forthcoming). Let  $X$  be some credence-forming process: that is, a process that outputs credences in a range of propositions. We can use  $R$  to score all of the credences that  $X$  produces. Average all of these scores, and we have a measure of  $X$ 's degree of reliability. Process reliabilists can then use this measure of reliability to give an account of justification for credences: a credence is (*prima facie*) justified iff it is produced by a reliable credence-forming process. What's the most suitable scoring rule for process reliabilism to use? Recent work has begun to tackle this question. In what follows, we confine ourselves to discussing two particularly prominent scoring rules—the Brier score and a calibration score.

Given its prominence in the literature, the Brier score is a natural option. But using the Brier score to measure the reliability of credence-forming processes faces challenges. For example, Dunn (2015) and Tang (forthcoming) object that if the Brier score is used, a credence-forming process that only outputs mid-level credences (say, a credence of .6) will never qualify as highly reliable; hence the credences it produces will never count as highly justified. Both Dunn and Tang object to this consequence. For instance, Tang argues that sometimes a particular input requires having a mid-level credence. If I have a vague visual experience of the silhouette of a horse, then it seems I should only have a mid-level credence that there is a horse in front of me: a credence of .6 in this proposition might well be justified, whereas a credence of 1 or 0 would not.

Another option is to measure the reliability of credence-forming processes using a *calibration score*. To see what it means for a credal state to be well-calibrated, consider the following example from van Fraassen:

Consider a weather forecaster who says in the morning that the probability of rain equals .8. That day it either rains or does not. How good a forecaster is he? Clearly to evaluate him we must look at his performance over a longer period of time. Calibration is a measure of agreement between judgments and actual frequencies... This forecaster was perfectly calibrated over the past year, for example, if, for every number  $r$ , the proportion of rainy days among those days on which he announced probability  $r$  for rain, equaled  $r$ . (van Fraassen 1984: 245)

According to the calibration approach, a credence is justified iff it's produced by a well-calibrated process. This avoids the objection to using the Brier score: after all, a credence-forming process that produces mid-level credences can still be well-calibrated.

However, the calibration approach has also elicited criticism. Goldman (1986) asks us to imagine an agent  $A$ , 70% of whose opinions turn out to be true.  $A$  can achieve a perfectly calibrated credence function by adopting a .7 credence in every proposition about which she has an opinion. However, Goldman argues that it's wrong to automatically conclude that  $A$ 's credal state is perfectly reliable: if  $A$  has no good reason for adopting a .7 credence in many of the propositions in question, then her credal state shouldn't count as justified. Dunn (2015) defends the calibration approach, arguing that the relevant question is whether the

*process* that produced  $A$ 's credal state is reliable. In order to answer this question, it's not enough to look at the truth-ratio of  $A$ 's opinions at the actual world; rather, we should look across a range of nearby worlds. If it's just a matter of chance that 70% of the propositions  $A$  has an opinion about are true, then by looking at the truth-values of  $A$ 's opinions at nearby worlds the calibration approach will be able to avoid the counterintuitive consequence that  $A$ 's credal state is perfectly reliable.

Weng Hong Tang (forthcoming) objects to the calibration approach on the grounds that a credence-forming process can be well-calibrated even though that process is insensitive to relevant evidence. In light of the perceived shortcomings of the calibration approach (and those of alternative scoring rules), Tang proposes a synthesis of reliabilism and evidentialism, where evidentialism can roughly be understood as the view that a belief's justification is determined by how well it is supported by the evidence that the believer has. According to Tang's proposal, a credence of  $C(p)$  is only justified if it is based on some ground  $g$ , such that the objective probability of the credence having a true content given  $g$  approximates  $C(p)$ . (Other syntheses of reliabilism and evidentialism are discussed in [section 6.2](#).)

Only recently have philosophers started to systematically explore the possibility of using scoring rules to provide a reliabilist theory of credal justification. However, given its position at the intersection of traditional and formal epistemology, this will likely prove to be a rich and important area of ongoing research.

## 6. Cousins and Spinoffs of Process Reliabilism

A number of theories have “branched off” from process reliabilism, borrowing some key ideas but parting company with respect to others. This section discusses two such cousins of process reliabilism: virtue reliabilism and syntheses of reliabilism and evidentialism.

### 6.1 Virtue Reliabilism

As its label suggests, virtue reliabilism is a branch of virtue epistemology that emerged in the mid-1980s in the wake of process reliabilism and shares some significant features with it. In particular, one of its central theoretical notions, that of an epistemic competence, resembles that of a reliable belief-forming process type. And its notion of the exercise of an epistemic competence resembles that of a token of a reliable process. Leading proponents of virtue reliabilism include Ernest Sosa (1991, 2007, 2010, 2015), John Greco (1999, 2009, 2010) and Duncan Pritchard (2012). Here we will focus primarily on Sosa's version.

Most virtue reliabilists do not explicitly use the notion of a “reliable process”, preferring instead the notions of “competence”, “virtue”, “skill” (Sosa) or “ability” (Greco, Pritchard). How should we understand these notions? Sosa often characterizes competences in terms of dispositions, for instance:

A competence is a certain sort of disposition to succeed when you try. So, exercise of a competence involves aiming at a certain outcome. It is a competence in part because it is a disposition to succeed reliably enough when one makes such attempts... It is thus tied to a conditional of the form: if one tried to  $\phi$ , one would (likely enough) succeed. (2015: 96)

*Epistemic competences*—the sort of competence that is relevant to epistemology—are dispositions of a specific variety: dispositions to arrive at the truth.

One question that arises for such accounts of competence is how we are to understand the dispositions in question. Are they general dispositions of an agent to arrive at the truth about some matter? Or are they to be understood as implicitly relativized to belief-forming processes or methods, in which case an epistemic competence is really of the form: *a disposition to arrive at the truth when employing process P*?

From a process reliabilist perspective, it's necessary to relativize the dispositions to belief-forming processes or methods. After all, process reliabilists will insist that in order to know whether an agent's belief that  $p$  is justified or counts as knowledge, it's not enough to know whether the agent is generally

disposed to arrive at truths about  $p$ -related matters. Instead, we'll need to know whether the agent's particular belief that  $p$  was the result of a reliable process. (After all, an agent could form the belief that  $p$  via an ultra-reliable process, even if she's generally disposed to form beliefs about  $p$ -related matters in highly unreliable ways.) In effect, committed process reliabilists will suggest that virtue reliabilists face a dilemma: either epistemic competences are general dispositions of the agent, in which case they won't be able to perform the various jobs required of them (specifically, explaining whether a belief is justified, or amounts to knowledge), or they are implicitly relativized to processes, in which case epistemic competences are not significantly different from reliable belief-forming processes. In the latter case, epistemic competences "collapse" into reliable processes.

In at least some discussions of epistemic competences, virtue reliabilists indicate a willingness to relativize epistemic competences to processes. For example, Sosa describes good eyesight and color vision as paradigmatic epistemic competences (Sosa 1991: 271, 2010: 467)—both of which are also standard examples of reliable processes. Similarly, Greco (1999) suggests that the intellectual virtues are processes that form stable components of an individual's cognitive character—suggesting that they are a type of reliable process.

If epistemic competences are understood as involving reliable processes, then virtue reliabilism inherits many of the challenges facing process reliabilism—in particular, the generality problem. (In virtue reliabilist terms, this will amount to the question: "How exactly should we type epistemic competences?") Of course, this result is unsurprising if Comesaña is right that the generality problem is a problem for everyone who tries to give an adequate theory of justified belief.

Virtue reliabilism differs from traditional process reliabilism in its choice of analysandum. Historically process reliabilists have focused on giving an account of *justification*; by contrast, virtue reliabilists have focused on giving an account of *knowledge*. However, one certainly could try to extend one's virtue reliabilism to justification. Indeed, if one assumes that knowledge entails justification, being a virtue reliabilist about the former seems to lead naturally to virtue reliabilism about the latter. And if epistemic competences are understood as reliable processes, the resulting virtue reliabilist account of justification would presumably amount to a version of process reliabilism.

Let us turn now to virtue reliabilist accounts of knowledge. How do virtue reliabilists propose to understand knowledge in terms of epistemic competences? There are a variety of slightly different proposals in the literature (Greco 2009, 2010; Sosa 2007, 2015; Turri 2011). However, virtue reliabilists typically understand knowledge as involving some sort of explanatory relation between having a true belief and the exercise of an epistemic competence. For instance, Sosa (2007) holds that  $S$  knows  $p$  iff  $S$  aptly believes  $p$ , where  $S$ 's belief is apt iff it is correct *because* of the exercise of an epistemic competence (see also Greco 2009, 2010 for a closely related account). More recently, Sosa (2010, 2015) defends a similar account couched in terms of "manifestation": knowledge is belief whose correctness manifests the agent's epistemic competence (see Turri 2011 for a similar account).

How do such accounts handle Gettier cases? Sosa 2007: 94–97 discusses Lehrer's (1965) Nogot/Havit case, in which a subject  $S$  truly believes that someone here owns a Ford, but he only does so on the basis of Nogot's misleading testimony. Sosa claims that while  $S$  holds this belief because of the exercise of an epistemic competence,  $S$ 's belief isn't correct *because* of the exercise of an epistemic competence. This explanation raises important questions about how to understand the relevant "because of" relation here: what exactly is the difference between a true belief being held because of an epistemic competence, and a belief being correct because of an epistemic competence? (After all, the exercise of the competence virtually never plays any role in making true the *proposition* believed. This proposition is true because of what happens in the world.) Other ways of fleshing out the details of a virtue reliabilist analysis raise similar questions. (See Lackey 2007 for critical discussion of virtue reliabilist theories that appeal to a notion of "epistemic credit".)

Even if a virtue reliabilist account of knowledge can handle some Gettier cases, there remains a question of whether it will be able to handle the full spectrum of Gettier cases. One case that has been thought to cause particular trouble for virtue reliabilists is the fake barn scenario (discussed by Goldman 1976,

credited to Carl Ginet). In the fake barn scenario, Henry sees from the road the one genuine barn in an area filled with many convincing barn façades. Henry forms a true belief that there's a barn in front of him; what's more, the fact that he correctly believes there's a barn in front of him seems to be causally explained by an exercise of his visual competence.

In response, one option is simply to embrace the conclusion that Henry's belief amounts to knowledge (Sosa 2010). This strikes many to be counterintuitive: the "received" view on fake barn cases is that they are cases of non-knowledge. It is true that a recent experimental philosophy study found that lay people do attribute knowledge to protagonists in fake-barn cases, although they are less inclined to do so than in unproblematic cases of knowledge (Colaco et al. 2014). This result may lend support to Henry's knowing. However, if, as many philosophers claim, philosophers themselves have more expertise than laypersons do in judging such cases, the power of the indicated result is debatable. It is questionable how much this finding rescues virtue reliabilists from a potentially severe counterexample.

Another response is to abandon the hope that virtue reliabilism on its own will solve every Gettier case. Pritchard (2012) takes this line, opting for a view that combines elements of virtue epistemology with a safety requirement on knowledge (where, again, safety is roughly the requirement that the belief in question couldn't easily have been held falsely). On Pritchard's view, Henry's belief that he sees a barn is unsafe, hence fails to count as knowledge.

A full assessment of these issues is beyond the scope of the current article. (For further discussion of whether virtue reliabilism can handle Gettier cases, see Miracchi 2015.<sup>[2]</sup>) This much is clear: one feature that distinguishes virtue reliabilism from classical process reliabilism is its distinctive treatment of knowledge. However, this treatment gives rise to serious questions and challenges which have not yet been fully resolved.

## 6.2 Syntheses of Reliabilism and Evidentialism

Process reliabilism and evidentialism have long been viewed as competitors, even antitheses of one another, with one of them (reliabilism) being a paradigm of externalism and the other (evidentialism) a paradigm of internalism. However, Juan Comesaña (2010) and Alvin Goldman (2011), both reliabilists, have toyed with the prospect of combining the best features of each theory to form a new theory that evades earlier problems. From the perspective of this entry, the chief question is whether such an accommodation helps reliabilism with some of its problems while not abandoning its essential features.

Juan Comesaña (2010) suggests that a hybrid of reliabilism and evidentialism along the following lines can help evade some of the problems facing traditional reliabilism:

**Proto-evidentialist Reliabilism:** *S*'s belief that *P* is justified if and only if that belief was produced by a process *X* which includes some evidence *e* and:

1. *e* doesn't include any beliefs of *S* and *X* is actually reliable; or
2. *e* includes beliefs of *S*, all of these beliefs are justified, and *X* is conditionally actually reliable.

(Note that this is not Comesaña's final theory. His final proposal—"Evidentialist Reliabilism"—involves certain further features designed to help solve the generality problem. For simplicity, we focus on his more basic proposal for integrating evidentialism with reliabilism.)

There are two main respects in which this departs from more traditional versions of reliabilism. The most obvious is that it involves an *evidential requirement*. This is intended to help with the case of Norman the clairvoyant. One crucial feature of Norman's situation is that he has no evidence for or against his clairvoyance powers, or regarding the whereabouts of the President. This is at least one of the reasons (says Comesaña) why we have the intuition that Norman is not justified. By imposing an evidential requirement on justification, we seem to solve the Norman problem.

However, there remains the question of how to understand the relevant notion of "evidence". As we've

seen (§2), traditional process reliabilists resisted defining “justification” in terms of evidence because they didn’t want an analysis that relied on any unreduced epistemic notions (Goldman 1979). But even for those who do not share these reductive ambitions, it’s natural to want at least some account of the sort of evidence that Proto-Evidentialist Reliabilism invokes.

Comesaña suggests following Conee and Feldman (2004) in opting for a “mentalist” construal of our evidence, according to which our evidence ultimately consists in various mental states. For those who find this approach promising, there remains the difficult question of *which* mental states constitute a subject’s evidence. (Are they conscious experiences? States that are accessible to consciousness? Beliefs?) A fully worked out hybrid of reliabilism and evidentialism would hopefully include answers to these questions.

A less obvious departure from traditional forms of reliabilism is the lack of any overtly historical condition on justifiedness. Traditional forms of reliabilism make epistemic status of a belief at a time *t* depend not only on features of the agent at *t*, but also on facts about how the believer acquired the belief in question. Here’s an example that motivates this “historicist” dimension to traditional reliabilism (Goldman 1999). Last year Sally read about the health benefits of broccoli in a New York Times science section story. She then forms a justified belief in broccoli’s beneficial effects. She still retains that belief today but no longer recalls the evidence she had upon first reading the story. And she hasn’t encountered any further evidence in the interim, from any kind of source. Isn’t her belief in broccoli’s beneficial effects still justified? Presumably this is because of her past acquisition. True, she also has a different kind of evidence, namely, her (justified) belief that whenever she seems to remember a (putative) fact it is usually true. But this is not her entire evidence. It is an important determinant of her belief’s J-status at *t* that she was justified in forming it originally on the basis of good evidence (of another kind). Had her original belief been based on very poor evidence, e.g., reading a similar story in an untrustworthy news source, so that the belief wasn’t justified from the start, her belief at time *t* would be unjustified—or at least much *less justified*. This indicates that the evidence she acquired originally still has *some impact* on the J-status of her belief at *t*. It is also important, of course, that the central process she uses to retain her broccoli belief, namely, preservative memory, is (conditionally) reliable. Its output beliefs tend to be true if its input beliefs are true. In light of such considerations, some process reliabilists will be reluctant to abandon a “historicist” component.

Thus far we have discussed Comesaña’s rationale for a synthesis between reliabilism and evidentialism. What about Goldman’s rationale? One rationale, presented “up-front” in his article (Goldman 2011), is to advance a “two-dimensional” approach to justification, an approach that makes room for a “degree of support” dimension of justification as well as a “proper causal generation” dimension. Evidentialism is centered on the “degree of support” dimension (sometimes phrased in terms of the “fittingness” of someone’s taking some doxastic attitude toward a proposition to the evidence that the person has) and offers little with respect to the second dimension. Reliabilism does the reverse: it does well on the causal generation dimension but poorly on the degree of support dimension. Why not add something to reliabilism to repair this possible lacuna? One might respond that this maneuver would be radically out of keeping with reliabilism. Doesn’t reliabilism utterly reject the relevance of evidence and strength of evidence in its whole approach?

It is true that process reliabilism has traditionally made no explicit reference to evidence—it never uses this term in its central theoretical principles. Nonetheless, it is well-positioned to make up for this deficit without much awkwardness. Consider all of the mental or psychological *states* that reliabilism traditionally appeals to. In addition to the *processes* it discusses, there are also *states* that serve as *inputs* to those processes. This includes both doxastic states (beliefs, primarily) and various experiences (perceptual, memorial, etc.). Although reliabilists never *call* these states “evidence” (nor refer to their contents as “evidence”), there is no reason why this couldn’t be done.<sup>[3]</sup> This remains an open question. The same points may be made in the language of “reasons” rather than “evidence”. Although traditional process reliabilists don’t find a place for reasons in their story, it is easy for them to say that justified belief states that are inputs to their beloved processes constitute reasons for their holding the output beliefs. Their quality as reasons may depend on their relations of support vis-à-vis the propositions inferred (and believed). Here again the upshot would be a two-dimensional theory of justification, according to which



the J-status of a belief is determined by both (i) the reliability of the belief-forming process, and (ii) the support provided by the evidence/reasons that serve as inputs to this process. Goldman makes some suggestions about how this would go, but such a task remains to be done more thoroughly. And if one hopes to fuse the two dimensions into a single measure of justifiedness, this seems conceivable but far from straightforward.

## Bibliography

- Alston, William P., 1980, "Level-Confusions in Epistemology", *Midwest Studies in Philosophy*, 5: 135–150.
- , 1995, "How to Think about Reliability", *Philosophical Topics*, 23(1): 1–29.
- Armstrong, David M., 1973, *Belief, Truth and Knowledge*, Cambridge: Cambridge University Press.
- Ball, Brian and Michael Blome-Tillman, 2013, "Indexical Reliabilism and the New Evil Demon", *Erkenntnis*, 78(6): 1317–1336.
- Beddor, Bob, 2015, "Process Reliabilism's Troubles with Defeat", *The Philosophical Quarterly*, 65(259): 145–159.
- Beebe, James R., 2004, "The Generality Problem, Statistical Relevance and the Tri-Level Hypothesis", *Noûs*, 38(1): 177–195.
- Bishop, Michael., 2010, "Why the Generality Problem is Everybody's Problem", *Philosophical Studies*, 151: 285–298.
- BonJour, Laurence, 1980, "Externalist Theories of Knowledge", *Midwest Studies in Philosophy*, 5: 53–73.
- Brier, G., 1950, "Veritification of Forecasts Expressed in Terms of Probability", *Monthly Weather Review*, 78(1): 1–3.
- Chisholm, Roderick, 1966, *Theory of Knowledge*, Englewood Cliffs, NJ: Prentice Hall.
- Christensen, David, 2007, "Three questions about Leplin's Reliabilism", *Philosophical Studies*, 134(1): 43–50.
- Cohen, Stewart, 1984, "Justification and Truth", *Philosophical Studies*, 46: 279–296.
- , 2002, "Basic Knowledge and the Problem of Easy Knowledge", *Philosophy and Phenomenological Research*, 65(2): 309–329.
- Colaco, D., W. Buckwalter, S. Stich, and E. Machery, 2014, "Epistemic Intuitions in Fake-Barn Thought Experiments", *Episteme*, 11(2): 199–212.
- Comesaña, Juan, 2002, "The Diagonal and the Demon", *Philosophical Studies*, 110: 249–266.
- , 2006, "A Well-Founded Solution to the Generality Problem", *Philosophical Studies*, 129(1): 27–47.
- , 2010, "Evidentialist Reliabilism", *Noûs*, 94: 571–601.
- Conee, Earl and Richard Feldman, 1998, "The Generality Problem for Reliabilism", *Philosophical Studies*, 89(1): 1–29.
- , 2004, *Evidentialism: Essays in Epistemology*, Oxford: Oxford University Press.
- DeRose, Keith, 1992, "Contextualism and Knowledge Attributions", *Philosophy and Phenomenological Research*, 52(4): 913–929.
- , 1995, "Solving the Skeptical Problem", *Philosophical Review*, 104(1): 1–52.
- , 2009, *The Case for Contextualism: Knowledge, Skepticism, and Context* (Volume 1), Oxford: Clarendon Press.
- Douven, Igor and Christoph Kelp, 2013, "Proper Bootstrapping", *Synthese*, 190(1): 171–185.
- Dretske, Fred, 1970, "Epistemic Operators", *The Journal of Philosophy*, 67: 1007–1023.
- , 1971, "Conclusive Reasons", *Australasian Journal of Philosophy*, 49: 1–22.
- , 1981, *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.
- Dunn, Jeff, 2015, "Reliability for Degrees of Belief", *Philosophical Studies*, 172(7): 1929–1952.
- Feldman, Richard, 1985, "Reliability and Justification", *The Monist*, 68(2): 159–174.
- , 2003, *Epistemology*, Upper Saddle River, NJ: Prentice-Hall.
- Feldman, Richard and Earl Conee, 2001, "Internalism Defended", *American Philosophical Quarterly*, 38(1): 1–18.
- Foley, Richard, 1985, "What's Wrong with Reliabilism?" *The Monist*, 68: 188–202.
- Fricker, Elizabeth, forthcoming, "Unreliable Testimony", in Kornblith and McLaughlin forthcoming.
- Gilbert, Margaret, 1989, *On Social Facts*, London: Routledge.

- Goldberg, Sanford, 2010, *Relying on Others: An Essay in Epistemology*, Oxford: Oxford University Press.
- Goldman, Alvin I., 1975, "Innate Knowledge", in Stephen P. Stich (ed.), *Innate Ideas*, Berkeley: University of California Press, pp. 111–120.
- , 1976, "Discrimination and Perceptual Knowledge", *The Journal of Philosophy*, 73: 771–791.
- , 1979, "What Is Justified Belief?" in G.S. Pappas (ed.), *Justification and Knowledge*, Dordrecht: Reidel, pp. 1–25; reprinted in A.I. Goldman, *Reliabilism and Contemporary Epistemology*, New York: Oxford University Press, 2012, pp. 29–49.
- , 1986, *Epistemology and Cognition*, Cambridge, MA: Harvard University Press.
- , 1992, "Epistemic Folkways and Scientific Epistemology", in *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, Cambridge, MA: MIT Press, pp. 155–175.
- , 1999, "Internalism Exposed", *The Journal of Philosophy*, 96: 271–293.
- , 2008, "Immediate Justification and Process Reliabilism", in Q. Smith (ed.), *Epistemology: New Essays*, New York: Oxford University Press, pp. 63–82.
- , 2009, "Internalism, Externalism, and the Architecture of Justification", *The Journal of Philosophy*, 106: 309–338.
- , 2011, "Toward a Synthesis of Reliabilism and Evidentialism", in T. Dougherty (ed.), *Evidentialism and Its Discontents*, New York: Oxford University Press, pp. 254–290.
- , 2014, "Social Process Reliabilism: Solving Justification Problems in Collective Epistemology", in J. Lackey (ed.), *Essays in Collective Epistemology*, Oxford: Oxford University Press, pp. 11–41.
- Graham, Peter J., 2012, "Epistemic Entitlement", *Noûs*, 46(3): 449–483.
- Greco, John, 1999, "Agent Reliabilism", in J. Tomberlin (ed.), *Philosophical Perspectives* (Volume 13: Epistemology), Atascadero, CA: Ridgeview Press, pp. 273–296.
- , 2009, "Knowledge and Success from Ability", *Philosophical Studies*, 142: 17–26.
- , 2010, *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*, Cambridge: Cambridge University Press.
- Jones, W., 1997, "Why Do We Value Knowledge?" *American Philosophical Quarterly*, 34: 423–440.
- Jönsson, Martin, 2013, "A Reliabilism Built on Cognitive Convergence: An Empirically Grounded Solution to the Generality Problem", *Episteme*, 10(3): 241–268.
- Kagan, S., 1992, "The Limits of Well-being", *Social Philosophy and Policy*, 9(2): 169–189.
- Kornblith, Hilary, 2009, "A Reliabilist Solution to the Problem of Promiscuous Bootstrapping", *Analysis*, 69: 263–267.
- Kornblith, Hilary and B. McLaughlin (eds.), forthcoming, *Alvin Goldman and His Critics*, Oxford: Blackwell.
- Kvanvig, Jonathan, 2003, *The Value of Knowledge and the Pursuit of Understanding*, Cambridge: Cambridge University Press.
- Lackey, Jennifer, 2007, "Why We Don't Deserve Credit for Everything We Know", *Synthese*, 158: 345–361.
- Lehrer, Keith, 1965, "Knowledge, Truth, and Evidence", *Analysis*, 25: 168–175.
- , 1990, *Theory of Knowledge*, Boulder, CO: Westview Press.
- Leplin, Jarrett, 2007, "In Defense of Reliabilism", *Philosophical Studies*, 134(1): 31–42.
- , 2009, *A Theory of Epistemic Justification*, Springer.
- List, Christian and Philip Pettit, 2011, *Group Agency: The Possibility, Design, and Status of Corporate Agents*, Oxford: Oxford University Press.
- Lyons, Jack, 2009, *Perception and Basic Beliefs: Zombies, Modules, and the Problem of the External World*, Oxford: Oxford University Press.
- , 2011, "Precis of Perception and Basic Beliefs", *Philosophical Studies*, 153: 443–446.
- MacFarlane, John, 2005, "The Assessment Sensitivity of Knowledge Attributions", *Oxford Studies in Epistemology*, 1: 197–233.
- Millikan, Ruth, 1984, *Language, Thought, and Other Biological Categories*, Cambridge, MA: MIT Press.
- Miracchi, Lisa, 2015, "Competence to Know", *Philosophical Studies*, 172(1): 29–56.
- Nozick, Robert, 1981, *Philosophical Explanations*, Cambridge, MA: Harvard University Press.
- Olsson, Erik, forthcoming, "A Naturalistic Approach to the Generality Problem", in Kornblith and McLaughlin forthcoming.
- Plantinga, Alvin, 1993, *Warrant: The Current Debate*, New York: Oxford University Press.
- Pollock, John, 1984, "Reliability and Justified Belief", *Canadian Journal of Philosophy*, 14: 103–114.


- Pritchard, Duncan, 2005, *Epistemic Luck*, Oxford: Oxford University Press.
- , 2012, “Anti-Luck Virtue Epistemology”, *Journal of Philosophy*, 109(3): 247–279.
- Rabinowicz, Wlodek and Toni Roennow-Rasmussen, 1999, “A Distinction in Value: Intrinsic and For Its Own Sake”, *Proceedings of the Aristotelian Society*, 100(1): 33–49.
- , 2003, “Tropic of Value”, *Philosophy and Phenomenological Research*, 66: 389–403.
- Ramsey, F.P., 1931, “Knowledge”, in R.B. Braithwaite (ed.), *Foundations of Mathematics and Other Logical Essays*, London: Kegan Paul, pp. 126–128.
- Riggs, Wayne D., 2002, “Reliability and the Value of Knowledge”, *Philosophy and Phenomenological Research*, 64: 79–96.
- Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Braem, 1976, “Basic Objects in Natural Categories”, *Cognitive Psychology*, 8: 383–439.
- Schmitt, Frederick, 1994, “The Justification of Group Beliefs”, in F.F. Schmitt (ed.), *Socializing Epistemology*, Lanham, MD: Rowman and Littlefield, pp. 257–287.
- , 2014, *Hume’s Epistemology in the Treatise: A Veritistic Interpretation*, Oxford: Oxford University Press.
- Sosa, Ernest, 1991, *Knowledge in Perspective: Selected Essays in Epistemology*, Cambridge: Cambridge University Press.
- , 1993, “Proper Functionalism and Virtue Epistemology”, *Noûs*, 27 (1): 51–65.
- , 1996, “Postscript to ‘Proper Functionalism and Virtue Epistemology’”, in J.L. Kvanvig (ed.), *Warrant in Contemporary Epistemology*, Lanham, MD: Rowman and Littlefield, pp. 271–280.
- , 2000, “Skepticism and Contextualism”, *Philosophical Issues*, 10: 1–18.
- , 2001, “Goldman’s Reliabilism and Virtue Epistemology”, *Philosophical Topics*, 29 (1/2): 383–400.
- , 2007, *A Virtue Epistemology*, Oxford: Clarendon Press.
- , 2010, “How Competence Matters in Epistemology”, *Philosophical Perspectives*, 24: 465–476.
- , 2015, *Judgment and Agency*, Oxford: Oxford University Press.
- Stalnaker, Robert, 1999, *Context and Content*, Oxford: Oxford University Press.
- Stine, Gail, 1976, “Skepticism, Relevant Alternatives, and Deductive Closure”, *Philosophical Studies*, 29: 249–261.
- Swinburne, Richard, 1999, *Providence and the Problem of Evil*, Oxford: Oxford University Press.
- Tang, Weng Hong, forthcoming, “Reliability Theories of Justified Credence”, *Mind*.
- Titelbaum, Michael, 2010, “Tell Me You Love Me: Bootstrapping, Externalism, and No Lose Epistemology”, *Philosophical Studies*, 149(1): 119–134.
- Turri, John, 2011, “Manifest Failure: The Gettier Problem Solved”, *Philosophers’ Imprint*, 11(8): 1–11.  
URL = <<http://hdl.handle.net/2027/spo.3521354.0011.008>>
- Unger, Peter, 1968, “An Analysis of Factual Knowledge”, *The Journal of Philosophy*, 65: 157–170.
- Van Cleve, James, 2003, “Is Knowledge Easy—or Impossible? Externalism as the Only Alternative to Skepticism”, in S. Luper (ed.), *The Sceptics*, Aldershot: Ashgate, pp. 45–59.
- Van Fraassen, Bas, 1984, “Belief and the Will”, *Journal of Philosophy*, 81: 235–256.
- Vogel, Jonathan, 2000, “Reliabilism Leveled”, *The Journal of Philosophy*, 97(1): 602–623.
- Weisberg, Jonathan, 2010, “Bootstrapping in General”, *Philosophy and Phenomenological Research*, 81(3): 525–548.
- , 2012, “The Bootstrapping Problem”, *Philosophy Compass*, 7(9): 597–610.
- Williams, Michael, forthcoming, “Internalism, Reliabilism and Deontology”, in Kornblith and McLaughlin forthcoming.
- Williamson, Timothy, 2000, *Knowledge and Its Limits*, Oxford: Oxford University Press.
- Wright, Larry, 1973, “Functions”, *The Philosophical Review*, 82: 139–168.
- Zagzebski, Linda, 1996, *Virtues of the Mind*, Cambridge: Cambridge University Press.
- , 2003, “The Search for the Source of the Epistemic Good”, *Metaphilosophy*, 34: 12–28.

## Academic Tools

 [How to cite this entry.](#)

 [Preview the PDF version of this entry](#) at the [Friends of the SEP Society](#).

 [Look up this entry topic](#) at the [Indiana Philosophy Ontology Project](#) (InPhO).

 [Enhanced bibliography for this entry](#) at [PhilPapers](#), with links to its database.

## Other Internet Resources

[Please contact the author with suggestions.]

## Related Entries

[contextualism, epistemic](#) | [epistemology](#) | [epistemology: naturalism in](#) | [epistemology: social](#) | [epistemology: virtue](#) | [justification, epistemic: foundationalist theories of](#) | [justification, epistemic: internalist vs. externalist conceptions of](#) | [knowledge: analysis of](#)

Copyright © 2015 by

[Alvin Goldman](#) <[goldman@philosophy.rutgers.edu](mailto:goldman@philosophy.rutgers.edu)>

[Bob Beddor](#) <[rbeddor@gmail.com](mailto:rbeddor@gmail.com)>

[Open access to the SEP is made possible by a world-wide funding initiative.](#)  
[Please Read How You Can Help Keep the Encyclopedia Free](#)

**Stanford** | Center for the Study of  
Language and Information

The Stanford Encyclopedia of Philosophy is [copyright © 2015](#) by [The Metaphysics Research Lab](#), Center for the Study of Language and Information (CSLI), Stanford University

Library of Congress Catalog Data: ISSN 1095-5054