Reliabilism

Olsson, Erik J

Published in:
Oxford Bibliographies Online

DOI:
10.1093/obo/9780195396577-0309

2016

Document Version:
Early version, also known as pre-print

Link to publication

Citation for published version (APA):

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
INTRODUCTION

On one characterization of “reliabilism” the term refers to any epistemological theory that emphasizes the reliability of some relevant epistemological feature. Thus Alvin Goldman explains reliabilism as a general approach to epistemology that emphasizes the truth-conduciveness of a belief-forming process, method, or other epistemologically relevant factor. On this broad characterization, reliabilist theories include those originally suggested under different labels, such as tracking theories. A tracking theorist holds that knowledge requires that the belief in question tracks truth. What this means is generally spelled out in counterfactual terms, such as “if the proposition p weren’t true, then the subject S wouldn’t believe that p”. On a narrower understanding, “reliabilism” denotes an epistemological theory that places special emphasis on the reliability or truth-conduciveness of belief forming.
processes, that is, processes leading up to the formation of a belief. According to Laurence Bonjour (2002, p. 244), “the central idea of reliabilism is that what makes a belief epistemically justified is the cognitive reliability of the causal process via which it was produced”. This is the view – commonly referred to as “process reliabilism” – that knowledge or justification is to be defined in terms of a reliable process of belief formation. Unless otherwise indicated, “reliabilism” will henceforward refer to process reliabilism. The basic idea of reliabilism, as a theory of knowledge, is that a subject S knows that p if and only if S’s true belief that p was obtained through a reliable process. Reliabilism, as a theory of justification, states that S’s belief that p is justified if and only if it was obtained through a reliable process. Thus, it is possible to view the reliabilist theory of knowledge as a variant of the traditional justified-true-belief account of knowledge. Reliabilists uniformly deny that the subject needs to have insight into, or awareness of, the existence of the reliable process leading up to her belief that p for her to know, or be justified in believing, that p. Facts of reliability may very well be external to the subject. For this reason, reliabilism is classified as an externalist, as opposed to an internalist, theory. Among the noted rationales for reliabilism we find the prospects of: 1) ascribing knowledge to higher animals, small children, certain artificial devices and so on; 2) providing a naturalized epistemology in which normative vocabulary has been eliminated; 3) solving some long-standing problems of epistemology, such as skepticism, the regress problem and the Gettier problem; and 4) providing answers to some otherwise puzzling problems, e.g. how one can know something without knowing that one does. As for the regress, it arises on the assumption that a person can know that p only if she is justified in believing that p. This is so because whatever other belief justifies the first must itself be justified, and so on. Since reliabilism does not assume that knowledge requires having a justified belief, the regress problem does not arise.

GENERAL OVERVIEWS AND TEXTBOOKS

Goldman’s entry on reliabilism in the Stanford Encyclopedia of Philosophy is an authoritative and up-to-date overview over the history and problems of reliabilism. There are a number of book length treatises that serve well as overviews. Lemos 2007 is a textbook which positions reliabilism in the context of modern epistemology. Goldman and McGrath 2014 is a state-of-the-art textbook on epistemology with a strong reliabilist focus.


This is an extensive and in-depth overview of reliabilism. Having introduced the reliabilist theory of knowledge, Goldman proceeds to discuss reliabilism as a theory of justification. He gives a detailed account of the standard problems and analyzes various modifications and refinements that have been proposed in response to the objections.

Lemos’s textbook is a useful introduction to the theory of knowledge generally and to reliabilism specifically. It contains several chapters devoted to reliabilist issues: on reliabilism and virtue epistemology, on internalism, externalism and epistemic circularity, and on naturalized epistemology.


Goldman and McGrath survey many of the core issues in contemporary epistemology, such as the structure of justification, defining knowledge and skepticism about knowledge. The book also takes up some new themes including social and probabilistic epistemology. The focus of the book is on naturalized epistemology in general and reliabilism in particular.

ANTHOLOGIES

Many of the central texts on reliabilism can be found in some excellent anthologies. One particularly useful collection is Bernecker and Dretske 2000 which contains a number of the most distinguished contributions to the topic. As for Goldman’s work on reliabilism, Goldman 2012 is recommended. Kornblith 1994 features some of the most important texts on naturalized epistemology. Schurz and Werning 2009 and Kornblith and McLaughlin forthcoming are two up-to-date anthologies on Goldman’s epistemology.


This anthology covers a range of classical papers on naturalized epistemology, by authors such as W.V.O. Quine, Alvin I. Goldman, Jaegwon Kim, Philip Kitcher and Stephen Stich. Topics range from justification and a priori knowledge to issues in logic and cognition.


Bernecker and Dretske’s anthology is a highly recommendable collection of outstanding work in epistemology featuring many reliabilist classics, including seminal texts by Armstrong, Dretske and Goldman. The texts are thematically organized and each group of papers is preceded by a highly readable introduction and overview written by the Editors.


In this volume, Goldman has assembled his most important writings on reliabilism and related subjects. The book starts with Goldman’s early work on justified belief which is followed by articles addressing topics like internalism and externalism, reliabiilism and evidentialism, and the value of knowledge. The concluding articles deal with issues of social epistemology and epistemological methodology.

This anthology on contains original articles discussing various aspects of Goldman’s reliabilism. Goldman replies to his critics in a final chapter.


This is an extensive volume featuring original articles on most aspects of Goldman’s epistemology, including his reliabilism, written by leading theorists of knowledge. The volume ends with Goldman replying to his critics.

**HISTORICAL BACKGROUND**

The historical background can be divided into *early defenses* and *contemporary defenses*

**Early defenses**

F. P. Ramsey is generally credited with having provided the first formulation of the reliabilist theory of knowledge. In a note, reproduced in Ramsey 1931, he presents what may be the first clear formulation of a reliabilist theory of knowledge writing that a belief counts as knowledge if was true, certain, obtained by a reliable process. Still, he thinks that the third condition is in need of some amendment. Ramsey also considered the application of reliabilism to beliefs arrived at through inference. He pointed out that unreliable, as predicated of a process, seems to refer only to a fallacious method and not to a false premise. If, in a process of inference, occur other intermediary beliefs they must all be true ones. In this connection, Ramsey proposed to rephrase the third, reliability, condition: the belief in question must be formed in a reliable way. This note of Ramsey’s attracted little attention at the time. Instead, the theory, in its broader sense, was revived four decades later by a number of researchers. About 40 years after Ramsey a number of researchers proposed epistemological theories where the concept of reliability plays an essential role. They include Goldman 1967 and 1976, Dretske 1971 and Armstrong 1973. The truth-tracking theory introduced in Nozick 1981 also belongs to this category. The first statement of a reliabilist theory of justification is likely to be that offered in Goldman 1979. Kitcher 1980 defends a reliabilist theory of a priori knowledge and Goldman 1986 a reliabilist theory of knowledge in Ramsey’s spirit.


This is a collection of various texts and notes by Ramsey, whose seminal reliabilist definition of knowledge is found on p. 258.


Goldman presents a causal theory of knowledge which is acknowledged as a precursor to the more popular reliabilist theory. The causal theory is introduced in response to the Gettier problem, i.e., roughly, the problem of accounting for cases of justified true belief in which the belief in question is true by chance.

Dreske defends the view that knowledge requires conclusive reasons, thus emphasizing the reliability of reasons for a belief. The main idea is that the required kind of reasons be such that, given the reasons, the possibility that the belief is false can be eliminated.


Armstrong formulates what he refers to as the thermometer model of knowledge. A thermometer reading is a reliable indicator of the actual temperature. In the same spirit, a belief that is knowledge is a reliable indicator of the truth of the belief.


Goldman develops what has been called a relevant-alternatives model of knowledge. According to this view, knowledge requires that the person holding the belief is able reliably to discriminate the truth of the belief from relevant alternatives. The theory is developed partly in response to objections raised against the causal theory stated in Goldman 1976.


Goldman formulates what is taken to be the first reliabilist account of justification. A belief is justified if it was acquired through a reliable process of belief formation, i.e. one that tends to produce true beliefs in a majority of cases. A reliable inference process confers justification on a belief only if its input beliefs are themselves justified by reliable processes. Goldman also adds a non-defeater condition: for a belief to be justified the person must not have evidence against the reliability of the process she employed. This requirement, too, is spelled out in reliabilist terms.


Kitcher defends a reliabilist theory of a priori knowledge. The proposal is that a person knows a priori that p just in case she has a true belief and that belief was produced by a process which is an a priori warrant for it. A priori warrants are moreover assumed to be ultra-reliable: in a counterfactual situation in which an a priori warrant produces the same belief that belief is true.


Nozick introduces a counterfactual theory according to which the characteristic feature of knowledge is that it tracks truth: if the belief were true, the subject would (still) believe it; and if it were false, the subject would not believe it. Nozick goes on to argue that the method (or way of believing) should be held fixed in the analysis. For that purpose he redefines knowledge as a three-place relation: a subject S knows that p through method (or way of believing) M.
In this seminal book, Goldman details a reliabilist theory of knowledge and justification. A characteristic feature is a distinction between the global and local reliability of a process. The former is the reliability for all (or many) uses of the process. The latter is the reliability of the process with respect to a particular belief. Knowledge requires, according to Goldman, a process that is reliable in both senses. The book contains extensive discussions of most issues of reliabilism often drawing on work in cognitive science. Goldman considers not just the reliability of belief forming processes but also their power and speed.

Contemporary defenses

A number of contemporary authors defend reliabilism, mostly as a theory of knowledge rather than of justification. Usually these authors defend a modified version of the original theory, e.g. a version of reliabilism which emphasizes social factors, in contrast to the individualistic picture that characterized early defenses of reliabilism. Apart from the recent books by Goldman already mentioned two notable general book length defenses of reliabilism are Kornblith 2002 and Goldberg 2010. Papinau 1992 is a shorter defense of classic reliabilism focusing on induction and skepticism. A number of authors have advanced modified versions of reliabilism, e.g. theories that combine reliability with elements from other epistemological theories. They include Brandom 1998, Greco 1999, Steup 2004 and Comesâña 2010.


Kornblith’s book is a novel defense of the reliabilist theory of knowledge from the point of view of particular species of naturalized epistemology, one according to which knowledge should be studied as a natural kind. This is in contrast to the methodology of conceptual analysis that is predominant in epistemology. Kornblith argues that the reliabilist conception is the one most congenial to this picture.


Goldberg considers knowledge from testimony as a challenge to an individualistic reliabilist theory of knowledge which holds that the process through which beliefs are formed never extend beyond the boundaries of the individual. Goldberg develops an alternative social-reliabilist approach according to which the relevant processes can involve reliance on others.


Papineau defends reliabilism as a solution to traditional skepticism about induction. He observes that the mere fact that inductive reasoning may occasionally lead to a false conclusion does not show that induction is not reliable. To the extent that our inductive practices are reliable, they can lead to knowledge. Papineau believes that similar considerations can undermine more ambitious forms of skepticism.

Brandom defends an inferentialist interpretation of reliabilism according to which we should understand assessments of reliability as issues of what inferences to endorse. This approach, he maintains, avoids the generality problem (see below) by showing that the different process types correspond to different (true) assumptions or auxiliary hypotheses that can be used as premises in one’s inferences.


Greco holds, first, that in order to avoid skepticism we must subscribe to some form of reliabilism and, second, that reliabilism, as standardly conceived, needs to be reformulated so as to disqualify “strange and fleeting” processes of belief formation. Greco concludes that what he calls agent reliabilism avoids skepticism without succumbing to the problem just mentioned.


Steup introduces and defends an epistemological theory which combines reliabilist and internalist elements. In particular, Steup defends the view that a subject’s sense experiences are a source of justification for her if and only if she has a memory impression of a track record of both perceptual and memorial success.


The author defends a hybrid account of epistemic justification which incorporated elements from both reliabilism and evidentialism (the view that a person is justified in holding a belief if that belief fits the evidence that the person has).

STANDARD OBJECTIONS

Early objections to reliabilism include the generality problem and various examples purporting to establish that reliability is not sufficient, or not necessary, for knowledge or justification. They include the new evil demon problem, clairvoyance and Truetemp examples, and the Gettier problem.

The generality problem

The generality problem, which was raised already in Goldman 1979, focuses on the claim that any process can be described at different levels of generality, i.e. broadly (e.g. a process of observation) or broadly (e.g. a process of observation in utter darkness). Depending on how the process is described it can come out as either reliable or unreliable. Since there is no fact of the matter how a process should best be described there is also no fact of the matter whether it was reliable or not. If this is true, reliabilism is an uninformative theory. Heller 1995 argued that the generality problem arises when the context is not taken into account. The most influential work on the generality problem is Conee and Feldman 1998, which stresses the severity of the generality problem also against Heller’s contextualist treatment. Goldman 2014 adopts a solution to the generality problem put forward in Jönsson 2013 and Olsson (forthcoming).

Heller offers a contextualist approach to the generality problem arguing that the term reliable is radically sensitive to the evaluator’s context. The context picks out a unique process type. Hence, once the context is taken into account, there is no problem of generality.


This article is considered to be the most forceful work pressing the generality problem as a main obstacle for a reliabilist theory. Conee and Feldman argue that people will agree on how a given process should be described only if the particular process type is salient in the conversational context and that this holds only if the process type has been explicitly mentioned. In contexts where this is not the case, there will be a generality problem.


Comesaña argues that reliabilism can overcome the generality problem if it incorporates elements form evidentialism, i.e. the view that knowledge has to be based on evidence in a certain way.


Olsson argues that the generality problem, as it is commonly formulated, reliabilism if the latter is viewed simply as providing an analysis or explication of knowledge. Olsson also considers a different formulation due to Conee and Feldman 1998 but contends that this particular version is in conflict with the basic level tradition in cognitive psychology. He also outlines an experimental setup for testing Conee and Feldman’s empirical assumptions.


Jönsson conducted a psychological experiment designed to test the empirical claims made in Conee and Feldman 1998. In the experiment subjects were asked to describe the process of belief formation operative in film clips showing someone coming to believe something. Jönsson found that most people agreed on how to describe a given process even if the process type had not been mentioned in the context.

The new evil demon, clairvoyance and Truetemp

A prominent challenge to the necessity of reliability for justification is the new evil demon problem raised in Cohen 1994 and Foley 1985. Throughout the years, Goldman has proposed several different solutions to this kind of problem, e.g. Goldman 1979, 1988 and 1992. The two most well-known objections to the sufficiency of reliabilism for justification or knowledge are the clairvoyance examples


Cohen considers a thought example in which an evil demon makes us to believe that we have reliably acquired beliefs when in fact our beliefs where unreliably formed. He argues that in such a scenario we would still be justified in our beliefs. Cohen concludes, against a reliabilist theory of justification, that reliable belief formation is not necessary for justified belief.


Foley discusses, among other things, the new evil demon objection to reliabilism. He concludes that reliability is not a significant criterion with can be used to distinguish epistemically valuable (justified) kinds of beliefs from other kinds of beliefs. Foley still thinks that having a true belief which is caused by a highly reliable cognitive process can be sufficient for knowledge.


This article contains the well-known clairvoyance examples purporting to show that a person, a clairvoyant in the example, can be in a situation where her true belief was reliably acquired and yet fail to live up to the rationality standards required for her to be justified in her belief.


Bach proposed to deal with BonJour’s clairvoyance examples by introducing a distinction between a belief being justified per se and a person being justified in holding that belief. Reliabilism captures the former, but not the latter, concept of justification. Bach goes on to defend reliabilism in the context of default reasoning.


Goldman makes a distinction between strong and weak justification to tackle various counter examples to a reliabilist theory of justification, including the new evil demon problem. Strong justifiedness requires de facto reliability and weak justifiedness imposes no such requirement. In the case of the demon-deceived cognizer, his beliefs can be described as lacking strong justifiedness but possessing weak justifiedness in the sense that he is blameless in entertaining them.


Lehrer gives a famous example featuring a person, Mr. Truetemp, who unbeknownst to himself has had a reliable thermometer implanted in his brain and, as an effect thereof, forms completely reliable beliefs.
about the temperature. Lehrer maintains that since Mr. Truetemp has no reason to think that his beliefs are true, he lacks knowledge.


Goldman introduces a two-stage theory of justification in order to tackle counter examples to reliabilism. At the first stage, a mental list of epistemic virtues and vices is created, corresponding to “good” and “bad” ways of forming beliefs. The second stage consists in applying these virtues and vices to particular examples, such as the demon-deceived cognizer.


Lyons argues that consideration of demon worlds pose no real threat to reliabilism. The reason, Lyons maintains, is that reliabilist is correct to hold that the inhabitants of such worlds have very few justified beliefs.

The Gettier problem

Solving the Gettier problem was one of the motivations for Goldman’s early causal theory of knowledge. Since then various Gettier type examples have been discussed in the reliabilist literature and led to theory development. A famous example is the barn façade introduced in Goldman 1976 as a counterexample to the causal analysis. In Goldman 1986, Gettier examples provide a motivation for distinguishing between global and local reliability. In Gettier cases, the subject uses a process that is globally but not locally reliable. Zagzebski 1994 stresses the inescapability of the Gettier problem for internalist and externalist theories alike. Olsson 2015 argues, however, that the Gettier problem is not a decisive objection to reliabilism if the latter is viewed as providing a Carnapian explication of knowledge.

Zagzebski, Linda. The Inescapability of Gettier Problems.

Zagzebski argues, among other things that so long as the reliable belief forming processes required for knowledge do not have to be perfectly reliable, there is always the possibility that there will be a Gettier problem. Gettier problems are in that sense inescapable.


Olsson proposes that epistemologists embrace Carnap’s method of explication. According to Carnap, a good explication, or rational reconstruction, of a pre-systematic concept like knowledge should be fruitful, exact and simple while being as similar as possible to the original. Olsson argues that the reliabilist account of knowledge satisfies these four criteria to a high degree even if it does not give the right result in Gettier cases.
RECENT DEBATES

The recent debates include controversies about the value of knowledge and bootstrapping or easy knowledge. Critics have also raised difficulties in connection with lottery cases.

The value of knowledge

Knowledge is widely considered to be more valuable than mere true belief. It is thought that an epistemological theory that does not have this consequence is unacceptable. Reliabilism has been the specific target of this kind of criticism. In particular, it has been argued, e.g. by Kvanvig and Zagzebski, that the fact that a true belief was obtained through a reliable process does not give that belief an additional boost of value. Goldman and Olsson 2009 explores various reliabilist solutions to the value problem. Others have giving up reliabilism in favor of some form of virtue epistemology. The main thought is that knowledge arises out of the exercise of an epistemic virtue and that such exercise has value in itself. Kvanvig 2003 is a good introduction to the value problem. Haddock et al. 2009 is a useful collection of original contributions to the topic.


Kvanvig surveys the history of the value problem, from Plato and onward, and argues that it particularly affects reliabilism. He then argues that it can be solved, in its traditional form, by his favored version of virtue epistemology. Kvanvig also thinks that the problem in its most general form – showing why knowledge is more valuable than its conceptual parts – does not admit of a plausible solution. Instead, he argues that understanding, and not knowledge, is what has distinctive value.


The paper contains Goldman’s original defense of reliabilism in the face of the value problem. The solution Goldman offers has two elements: type-instrumentalism and value autonomization. The former ascribes value to process types rather than to process tokens. The latter is a psychological account of value assignment. Olsson offers a third proposal, the conditional probability solution, based on the idea that reliabilist knowledge is indicative of having more true beliefs (and knowledge) in the future.


This is recent collection of papers on the value of knowledge. Most authors who have made an original contribution to the debate are featured in the volume. Apart from Goldman and Olsson 2009, the volume includes articles by Greco, Kvanvig, Pritchard and others.

Bootstrapping and easy knowledge
The problem of bootstrapping was introduced in Vogel 2000 and stressed in Cohen 2002. Vogel returned to the problem in his 2008 article. To make a long story short, the worry is that reliabilism seems to license a kind of inference that makes the attainment of a certain kind of higher-order knowledge or justification too easy. If a process is reliable, then a person could come to know, or be justified in, this by induction from individual applications of the process. Goldman 2008 argues, following Van Cleve 2003, that if bootstrapping or easy knowledge is disallowed, the only alternative is skepticism, which is a worse option. Kornblith 2009 is a more recent attempt to undermine bootstrapping objections to reliabilism.


Vogel introduces the most discussed bootstrapping example with Roxanne who, simply by repeatedly reading off the gas gauge of her car, under certain circumstances can attain knowledge that the gauge is reliable. In general bootstrapping according to Vogel makes allow one to promote many, if not all, of one’s beliefs that were formed by reliable processes into higher-order knowledge that those beliefs were formed by reliable processes. Vogel thinks this is a decisive objection to reliabilism.


Cohen invokes unreflective (animal) and reflective knowledge to solve the bootstrapping or easy knowledge problem. Bootstrapping is prevented by assuming that instances of unreflective knowledge cannot combine individually with self-knowledge to generate inferences. In particular, a person cannot combine her unreflective knowledge that there is something red before her with her knowledge that her belief is produced by color vision in order to infer that her color vision is working correctly.


Van Cleve argues that if bootstrapping is disallowed, then skepticism ensues. The latter, in his view, is an option worse than the former.


Vogel continues the analysis of epistemic bootstrapping initiated in his 2000 article. This paper discusses, among other things, the relation between reliabilism and bootstrapping and between justification and bootstrapping.


Kornblith argues that Van Cleve’s 2003 paper does not fully succeed in defending reliabilism against charges of illegitimate bootstrapping. Kornblith proposes what he takes to be a more compelling defense.
Lotteries

Lottery problems, finally, arise for reliabilism, in the simplest case, due to the observation that a reliabilist should be able reliably to conclude, and know, that a given losing ticket will lose in a sufficiently large lottery in the face of the overwhelming probability of that proposition. This is in contrast to the common intuition that this is not something one can know on statistical ground alone. A more sophisticated lottery discussion can be found in Adler 2005. Comesâna 2009, following Alston 1988, argues that reliabilism needs to be spelled out in terms of the evidence upon which the subject based her belief, and that this move takes care of lottery examples.


This is an extended discussion on how to understand the concept of reliability employed by reliabilists. The paper focuses on lottery cases and other statistical examples.


Comesâna argues that lottery problems for reliabilism can be overcome if the latter is redefined so that it takes into account the evidence upon which the subject bases her belief. The relevant kind of reliability should be seen as high conditional probability of the belief on the evidence.