

Reliabilism as Explicating Knowledge: A Sketch of an Account

Olsson, Erik J

Published in:

Epistemology: Contexts, Values, Disagreement

2012

Link to publication

Citation for published version (APA):

Olsson, E. J. (2012). Reliabilism as Explicating Knowledge: A Sketch of an Account. In C. Jäger, & W. Löffler (Eds.), Epistemology: Contexts, Values, Disagreement Ontos Verlag.

Total number of authors:

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.

Download date: 16. Dec. 2025

Reliabilism as Explicating Knowledge: A Sketch of an Account

Erik J. Olsson

Olsson, E. J. (2012). Reliabilism as Explicating Knowledge: A Sketch of an Account. In Jäger, C. & Löffler, W. (Eds.) Epistemology: Contexts, Values, Disagreement. Ontos Verlag.

1. Introduction

According to simple reliabilism, S knows that *p* if and only if (1) *p* is true, (2), S believes that *p*, and (3) S's belief that *p* was acquired through (sustained by) a reliable process. Henceforth "reliabilism" means "simple reliabilism". The first clear statement of reliabilism apparently stems from F. P. Ramsey (1931). Since Ramsey, many distinguished researchers have advocated reliabilism in one form or another, including David Armstrong (1973), Fred Dretske (1969, 1971) and Alvin I. Goldman (1979, 1986). For a recent overview of reliabilism and the issues surrounding it, see Goldman (2011). In this paper I will sketch a defense of reliabilism from a certain methodological perspective, namely that of the theory of explication due to the great Austrian philosopher Rudolf Carnap. I will defend reliabilism (a) as a highly promising explication of knowledge, and also (b) as an explication of knowledge that is likely to be better than any other standard proposal on the market, although the latter claim will continue to have the status of a mere plausible conjecture.

2. Reliabilism as an explication of knowledge

An explication is a kind of definition satisfying the following conditions (Carnap 1950, 7):

- 1. The explicatum [the thing that explicates] is to be similar to the explicandum [the thing that is explicated] in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.
- 2. The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an exact form, so as to introduce the explicatum into a well-connected system of scientific concepts.
- 3. The explicatum is to be a fruitful concept, that is, useful for the formulation of many [true] universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).
- 4. The explicatum should be as simple as possible; this means as simple as the more important requirements (1), (2) and (3) permit.

Thus, to take Carnap's example, in zoology the artificial concept *piscis* has come to replace the common sense concept of fish, although *piscis* is a narrower concept that excludes several kinds of

animal which were subsumed under the concept fish, e.g. whales and seals. The main reasons for the replacement was that zoologists found that *piscis* is a much more fruitful concept than fish in the sense that it allows for the formulation of a greater number of interesting general truths: that all *piscis* are cold-blooded, extract oxygen from water using gills, lay eggs, and so on. At the same time, the concept *piscis*, apart from being a precise and simple one, is sufficiently close to the common sense concept of fish to replace it at least in a scientific context. This seems to be a typical case of a successful scientific definition. I will proceed on the assumption that scientific definitions of ordinary concepts generally should be seen as explications.

Now the most coherent/unified view on philosophical definitions is arguably that such definitions are, in principle, no different from scientific definitions. This follows for instance if one thinks that philosophy is a branch of science or is at least closely related to science or some variation on that theme. Since scientific definitions are explications, it would follow that philosophical definitions are, or should be seen as, explications as well. In fact, not only Carnap but also Quine (1960) held this view. In particular, definitions of knowledge are, or should be seen as, explications. There is an alternative route to this conclusion over a weak form naturalized epistemology stating that epistemology, but perhaps not philosophy in general, is a branch of science. Quine went on, of course, to advocate the naturalization of epistemology (Quine, 1969) which is significant in this context because reliabilism, as we all know, is usually subsumed under that epistemological program (although researchers tend to have different ideas about what exactly the program involves). For a recent defense of the method of explication, see Maher (2007).

For reliabilist knowledge to be a good explication of the ordinary concept of knowledge the following must be satisfied to a high degree:

- 1. Reliabilist knowledge must be similar the ordinary concept of knowledge in such a way that, in most cases in which the ordinary concept has so far been used, reliabilist knowledge can be used; however, close similarity is not required, and considerable differences are permitted.
- 2. The characterization of reliabilist knowledge must be given in an exact form, so as to introduce the concept into a well-connected system of scientific concepts.
- 3. Reliabilist knowledge must be a fruitful concept, that is, useful for the formulation of many universal statements (empirical laws, since it is a non-logical concept).
- 4. Reliabilist knowledge should be a simple concept, e.g. not requiring a complicated definition. That these conditions should be satisfied to a higher degree for reliabilism to be a good explication follows directly from Carnap's account of explication.

We understand the claim that reliabilist knowledge is a good explication, and probably better than other standard accounts, to mean that it is competitive in that regard among a certain class of definitions, all of which assume that knowledge involves true belief plus something more, e.g. JTB. Thus we leave the non-standard TB account out of the picture in this paper. It follows that we need not discuss the exactness, simplicity and so on of the concepts of truth and belief; they are after all

common to all competing definitions. Rather, we can focus entirely on the characteristic third clause of the reliabilist explication: that the belief has been acquired through a reliable process. We will deal with the claims 1-4 in the order 4, 2, 3, 1, moving from what I take to be the least to the most contentious.

3. Simplicity

Is the third reliabilist clause simple? It states that S's belief that p was acquired through a reliable process. Yes, that does sound quite simple. Are any of the competing accounts of knowledge simpler, e.g. JTB? No, probably not. It would seem that the reliabilist account of knowledge is at least as simple as any competing account of knowledge, and much simpler than most.

But hold on a second: isn't the notorious generality problem precisely an argument against the simplicity of the reliabilist account? Reliabilism does not include a method for deciding how a belief formation process should be classified and for deciding whether the process was reliable or not; it is therefore "seriously incomplete" (Conee and Feldman 1998). While reliabilism appears to be a simple theory on first sight, once it is supplemented with such methods it is bound to be extremely complex!

Plausible as it may sound this line of reasoning is mistaken. An explication of knowledge states necessary and sufficient conditions for knowledge. Providing additional criteria for deciding, in concrete cases, whether those conditions are satisfied is not part of what an explication of knowledge needs to deliver. Thus, in order to qualify as an explication a JTB style theory does not have to include additional criteria for deciding, in concrete cases, whether a given belief is justified. Nor does a reliabilist style explication have to provide a method for deciding whether the process was reliable (which would involve providing a further method for deciding how the process should be classified). Hence, from the point of view of the current methodology Conee and Feldman's reasoning fails: the fact that a reliabilist account of knowledge fails to include a method for identifying the process type and for deciding on the reliability of the process does not make that account incomplete as an explication of knowledge. An explication of knowledge is complete once it provides necessary and sufficient conditions for knowledge, which does not mean of course that the proposal cannot have other flaws but that's irrelevant in the present context.

There is a second point to be made in this connection. Let us by a *supplemented explication of knowledge* mean a package consisting of an explication of knowledge stating necessary and sufficient conditions in the usual way plus additional criteria for deciding when those first conditions obtain. Reliabilism conceived as a supplemented explication will be more complex than reliabilism conceived as an (unsupplemented) explication. But the same is true *mutuatis mutandis* of JTB, evidentialism or any other competing accounts of knowledge. For instance, a JTB theory which supplies criteria for deciding when a given belief is justified will be more complex than a JTB theory in the usual sense.

Be that as it may, the focus in this paper is on reliabilism as providing an (unsupplemented) explication, as opposed to a supplemented explication, of knowledge.

The bottom line is that our conclusion still stands: reliabilism provides a strikingly simple explication of knowledge. We will consider a more interesting version of the generality problem in section 6, in connection with Carnap's first condition, that of similarity with the ordinary concept.

4. Exactness

We recall that "the characterization of reliabilist knowledge must be given in an exact form, so as to introduce the concept into a well-connected system of scientific concepts". Consider again the characteristic reliabilist clause stating that S's belief that p was acquired through a reliable process. The crucial concept here, obviously, is that of a reliable process. This concept is a concept of statistics: a random process is reliable with regard to an outcome type (in this case "truth"), if outcomes of that type occur more often than not, or often enough for the purposes at hand. Hence, the reliabilist explication of knowledge introduces the concept of knowledge into *the well-connected system of scientific concepts represented by statistical theory*. I am not aware of any competing account of knowledge that to a similar high extent satisfies the condition of introducing the concept defined into a well-connected system of scientific concepts. JTB for sure does not because 'justified belief' is not a concept that occurs in exact science.

There is, to be sure, some vagueness here regarding how reliable a process must be to be reliable *simpliciter*. However, the same kind of vagueness pertains to many applications of statistics to real life, and it does not make statistics less exact. The context determines how reliable something must be in order to be reliable *simpliciter*. Thus, a reliable method for disposing of nuclear waste is hopefully more reliable than a reliable method for disposing of garden waste. And of course, competing definitions of knowledge show the same kind of vagueness: how justified must a belief be in order to be justified *simpliciter*?

We must conclude that the reliabilist explication of knowledge satisfies Carnap's desideratum of exactness to considerable degree due to its close connection to statistical theory, and it seems unlikely that its competitors could do better in this regard. This seems particularly unlikely for JTB, evidentialism and virtue epistemology, which all rely on concepts ('justification', 'evidence' and 'virtue') that arguably do not occur in exact science.

5. Fruitfulness

"Reliabilist knowledge must be a fruitful concept, that is, useful for the formulation of many universal statements (empirical laws, since it is a non-logical concept)". On my reconstruction, challenges to the fruitfulness of the reliabilist concept of knowledge purport to show that such knowledge is no more

valuable than mere true belief: the universal inference from knowledge to surplus value is rejected. The famous swamping argument by Jonathan Kvanvig (2003), Linda Zagzebski (1996) and others attempts to accomplish just this. I have rejected this argument in several previous papers, arguing for the following:

- (A) If you have reliabilist knowledge of p, then you are more likely to have more true belief (and knowledge) in the future than if you have a mere true belief that p (Goldman and Olsson 2009).
- (B) If you have reliabilist knowledge of p, then your true belief that p is more likely to be stable than if you have a mere true belief that p (Olsson 2007, 2008)
- (C) As Williamson (2000) has observed, stability of true belief is positively correlated with practical success when acting over time (Olsson 2007, 2008).

I take these three claims to express true statistical laws relating reliabilist knowledge to cognitive and practical success. The laws are true in our world because our world satisfies a number of empirical regularities such as:

(Non-uniqueness) Once you encounter a problem of a certain type, you are likely to face other problems of the same type in the future.

(Cross-temporal access) A method that was used once is often available when similar problems arise in the future.

(Learning) A method that was unproblematically employed once will tend to be employed again on similar problems in the future.

(Generality) A method that is (un)reliable in one situation is likely to be (un)reliable in other similar situations in the future.

The regularities hold normally (but not always) and this is sufficient for the first statistical law to be true (i.e. for the relevant conditional probabilities to be in place).

Consider that statistical law: if you have reliabilist knowledge of p, then you are more likely to have more true belief (and knowledge) in the future than if you have a mere true belief that p. Let us imagine the following train of events:

- 1. S acquires the true belief that p through method M as a solution to problem P (assumption)
- 2. S is confronted with a problem of the same type as P in the future (non-uniqueness)
- 3. *S* still has access to method *M* (*cross-temporal access*)
- 4. S makes use of M again (learning)
- 5. S acquires a new true belief q through method M

Given the assumption of generality, this fortunate sequence of events is more likely, if *M* is reliable, than it is if *M* is unreliable due to the fact that the step from 4 to 5 is more likely in that case. The other steps are likely, and equally so, for reliable and unreliable methods. In other words, under the condition stated, the probability that S will acquire new true beliefs is greater conditional on reliabilist knowledge than it is conditional on mere true belief. For a recent exchange regarding the learning

condition, see Jäger (2011a, 2011b), Olsson and Jönsson (2011). For replies to other criticisms, see Olsson (2009) and Olsson (2011a).

Carnap's account of explication requires of a fruitful empirical concept that it be useful for the formulation of many universal empirical laws. There are at least three empirical laws that involve reliabilist knowledge, namely those just mentioned. I am not aware that the competing accounts of knowledge, JTB etc., allow for the formulation of (true) empirical laws. I conjecture, therefore, that the reliabilist conception of knowledge satisfies Carnap's desideratum of fruitfulness to a uniquely high degree.

6. Similarity to the ordinary concept

"Reliabilist knowledge must be similar the ordinary concept of knowledge in such a way that, in most cases in which the ordinary concept has so far been used, reliabilist knowledge can be used; however, close similarity is not required, and considerable differences are permitted." In the literature, there are two main challenges to the claim that reliabilist knowledge is sufficiently similar to the ordinary concept of knowledge: the generality problem and the Gettier problem.

To set the stage for the generality problem, as I take it to be most fruitfully and compellingly understood, consider the following train of thought:

- 1. Each concrete (token) process can be classified as belonging to a great many different types.
- 2. It is not obvious how to single out one of these types rather than another as *the* unique associated type of the process in question.
- 3. Furthermore, depending on what type is singled out as special, we may get different verdicts as regards the reliability of the process.
- 4. Hence, there is no fact of the matter whether a given process is reliable or not.
- 5. Hence, there is no fact of the matter whether a given person has reliabilist knowledge and not. As for the third premise, a concrete process may be categorized both as "reading The Economist" and as "reading" (regardless of what is being read). In the first case is it probably reliable, in the second case probably not.

Under what conditions would the generality problem be a threat to the claim that reliabilist knowledge is reasonably similar to the ordinary concept of knowledge? It would be if it would show that it is not true that, in most cases in which the ordinary concept of knowledge has so far been used, reliabilist knowledge can also be used. This is indeed what Earl Conee and Richard Feldman have argued as I reconstruct them (Conee and Feldman 1998, Feldman and Conee 2002). Here is their argument in a nutshell:

1. People will agree on process typing and reliability only if a particular process type is salient in the conversational context

- A particular process type is salient in the conversational context only if it has been explicitly mentioned
- 3. In many cases in which no type has been explicitly mentioned, we still agree on the corresponding knowledge claims
- 4. Hence, in many cases people will agree on the corresponding knowledge claims without agreeing on process typing and reliability
- 5. Hence: it is (arguably) *not* true that, in most cases in which the ordinary concept of knowledge has so far been used, reliabilist knowledge can be used.

It would follow that reliabilist knowledge is not sufficiently similar to the ordinary concept of knowledge and that it therefore fails to satisfy Carnap's first criterion on an explication of knowledge.

Unfortunately, Conee and Feldman's argument is empirically unsound. The second premise is false: it is not true that a particular process type is salient in the conversational context only if it has been explicitly mentioned. I have questioned that premise on theoretical grounds drawing on to the influential basic level tradition in cognitive psychology (Olsson 2011b, Olsson forthcoming). According to that tradition, our conceptual taxomomies are largely culturally inherited and for many taxomonies there is a salient "basic level", i.e. a level which is more naturally, and more often, used than other level. For instance, it is more natural to classify the thing on which I am now sitting as a "chair" than as, say, a "wooden structure", "piece of furniture", or a "designer chair", even though these descriptions can all be truthfully applied to the thing in question. Basic level categories stand out as being highly informative, yet economical. From membership in the chair category, many things can be (defeasibly) inferred: that you can sit on it, that it has four legs, and so on. You cannot infer those things if you classify the thing as a "piece of furniture". You can do it if you classify it as a designer chair. However, that would be a more complex, and therefore cognitively less economical, classification. Note that salience here has nothing to do with what has been mentioned or not in a conversational context. Similarly, I have proposed, categories like "seeing", "hearing", and so on are naturally viewed as salient basic level categories for belief forming processes. The claim that there are basic level categories for belief formation processes was recently empirically confirmed in an experimental study (Jönsson forthcoming a, forthcoming b). The study found, among other things, that people tend to agree on how to classify belief forming processes even when no classificatory type has been explicitly mentioned.

Let us move on to the Gettier problem, which can also be seen, or reconstructed, as an objection to the claim that the reliabilist concept of knowledge satisfies Carnap's criterion of similarity to the ordinary concept. Under what conditions would the Gettier problem be a threat to the claim that reliabilist knowledge satisfies this desideratum? It would be if it would show that it is not true that, in most cases in which the ordinary concept of knowledge has so far been used, the reliabilist concept of knowledge can be used. But clearly the Gettier problem does not show this: Gettier cases are too rare to threaten the claim that reliabilist knowledge can be substituted for ordinary knowledge in most

cases. They are too rare for that purpose because they involve the consecutive occurrence of two improbable events: a proposition (Brown owns a Ford) that is strongly supported by evidence turns out nonetheless to be false and yet by sheer luck (Brown is in Barcelona) the target proposition (Brown owns a Ford or is in Barcelona) comes out true anyway. Hence, the Gettier problem can never constitute a *fatal* challenge for a reliabilist explication of knowledge.

A critic might still want to entertain the thought that reliabilist knowledge cannot replace the ordinary concept in Gettier cases because (i) according to reliabilism the person in a Gettier scenario knows whereas (ii) according to the folk concept of knowledge, that person doesn't know. But both (i) and (ii) can be contested. As for (i), reliabilism by itself doesn't predict anything about Gettier scenarios; it gives a definite prediction only against the background of a specific categorization of the belief forming process involved. And who knows how the "folk" will categorize the belief forming processes in Gettier cases? For instance, will the folk think of Henry's belief that there is a barn over there, referring to Goldman's barn façade case, as having been acquired through a process of "seeing" or through, say, a process of "seeing in deceptive circumstances"? We don't know. There is room for empirical work here. As for (ii), experimental work has undermined the supposed stability of intuitions about knowledge in Gettier cases (Weinberg et al 2001, Swain et al 2008). Swain et al (2008) report that the barn case elicited among a majority of subjects the "wrong" response: that the person before the barn ("Susan" in their example) knows that she is looking at a barn.

Gettier cases are the whales of epistemology: just as whales are like fish in some respects and like mammals in others, so too Gettier cases are like knowledge in some respects and like ignorance in others. They are like knowledge e.g. because they involve a (usually) reliable method of belief formation. They are like ignorance e.g. because they involve elements of randomness and error. So, are Gettier cases more like knowledge or more like ignorance? Consider the corresponding zoological question: are whales more like fish or more like mammals? Zoologists didn't attempt to answer this question by consulting their zoological intuitions. Rather they considered the theoretical costs and benefits of various classificatory options. I suggest that it is time that we do the same in epistemology.

7. Conclusion

My starting point was the methodology of explication which has strong independent standing in analytic philosophy deriving from its association with two of its founding fathers: Rudolf Carnap and W. V. O. Quine. It turned out that (simple) reliabilism looks very attractive indeed from that methodological perspective. The reliabilist explication of knowledge is strikingly simple, exact and fruitful. Furthermore, here is no compelling reason to believe that the reliabilist concept of knowledge falls short of being sufficiently similar to the ordinary concept. Even longstanding issues like the generality problem, in its most compelling form, and the Gettier problem were unable to undermine that conclusion. Future work includes the reconsideration, from an explanationist perspective, of some

further objections to reliabilism, including the problem of easy knowledge (Cohen 2002) and possibly Bonjour's clairvoyance cases (Bonjour 1980), although the latter target the reliabilist theory of justification rather than the reliabilist theory of knowledge. The present article has focused on JTB as the main competitor to reliabilism. We need to look more closely at other competing accounts, such as evidentialism and virtue epistemology, and how they fare as explications of knowledge before any more definite conclusions can be drawn.

References

Armstrong, D. M. 1973: Belief, Truth and Knowledge, Cambridge: Cambridge University Press.

Bonjour, L. 1980: Externalist Theories of Empirical Knowledge, *Midwest Studies in Philosophy* 5(1), 53-74.

Carnap, R. 1950: Logical Foundations of Probability, Chicago: Chicago University Press.

Cohen, S. 2002: Basic Knowledge and the Problem of Easy Knowledge, *Philosophy and Phenomenological Research* 65(2), 309-329.

Conee, E. and Feldman, R. 1998: The Generality Problem for Reliabilism, *Philosophical Studies* 89, 1-29.

Dretske, F. 1969: Seeing and Knowing, Chicago: Chicago University Press.

Dretske, F. 1971: Conclusive Reasons, Australasian Journal of Philosophy 49, 1-22.

Feldman, R. 1985: Reliability and Justification, The Monist 68, 159-174.

Feldman, R., and Conee, E. 2002: Typing Problems, *Philosophy and Phenomenological Research* 65 (1), 98-105.

Goldman, A. I. 1979: What is Justified Belief?, in: Justification and Knowledge, ed. G. Pappas,

Dordrecht: D. Reidel, 1-23.

Goldman, A. I. 1986: Epistemology and Cognition, Cambridge MA: Harvard University Press.

Goldman, A. I.: Reliabilism, *The Stanford Encyclopedia of Philosophy* (Summer 2011 Edition), Eward N. Zalta (ed.), URL = http://plato.stanford.edu/entries/reliabilism/>.

Goldman, A.I. & Olsson, E.J. 2009: Reliabilism and the Value of Knowledge, in: *Epistemic Value*, ed.

A. Haddock, A. Millar and D. H. Pritchard, Oxford: Oxford University Press, 19-41.

Jäger, C. 2011a: Process Reliabilism and the Value Problem, *Theoria* 77(3), 201-213.

Jäger, C. 2011b: Reliability and Future True Belief: Reply to Olsson and Jönsson, *Theoria* 77(3), 223-237.

Jönsson, M.L. forthcoming a: A Reliabilism Built on Basic Levels.

Jönsson, M.L. forthcoming b: Linguistic Convergence on Abstract Verbs: A Challenge to the Standard Explanation of Basic Levels.

Kvanvig, J. L. 2003: *The Value of Knowledge and the Pursuit of Understanding*, Cambridge: Cambridge University Press.

Maher, P. 2007: Explication Defended, Studia Logica 86 (2), 331-341.

Olsson, E. J. 2007: Reliabilism, Stability, and the Value of Knowledge, *American Philosophical Quarterly* 44(4), 343-355.

Olsson, E.J. 2008: Knowledge, Truth, and Bullshit: Reflections on Frankfurt. *Midwest Studies in Philosophy* 32: 94-110.

Olsson, E.J. 2009: In Defense of the Conditional Probability Solution to the Swamping Problem, *Grazer Philosophische Studien* 79, 93-114.

Olsson, E.J. 2011a: Reply to Kvanvig on the Swamping Problem, Social Epistemology 25(2), 173-182.

Olsson, E.J. 2011b: The Generality Problem Naturalized, Festschrift for Erik Carlson. *Uppsala Philosophical Studies*, in press.

Olsson, E. J. forthcoming: A Naturalistic Approach to the Generality Problem, in: *Goldman and His Critics*, ed. H. Kornblith and B. Maclaughlin, Blackwell.

Olsson, E. J., and Jönsson, M.L. 2011: Kinds of Learning and the Likelihood of Future True Beliefs: Reply to Jäger on Reliabilism and the Value Problem, *Theoria* 77(3), 214-222.

Quine, W. V. O. 1960: Word and Object, The MIT Press.

Ramsey, F. P. 1931, *The Foundations of Mathematics and other Logical Essays*, ed. R. B. Braithwaite, London: Routledge and Kegan Paul.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. 1976: Basic Objects in Natural Categories, *Cognitive Psychology* 8, 382-439.

Swain, S., Alexander, J., and Weinberg, J. 2008: The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp, *Philosophy and Phenomenological Research* 126(1), 138-155.

Weinberg, J. M., Nichols, S., and Stich, S. 2001, "Normativity and Epistemic Intuitions", *Philosophical Topics* 29, 429-460;

Zagzebski, L. 1996: Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge, Cambridge: Cambridge University Press.