Philos Stud (2015) 172:57–72
DOI 10.1007/s11098-014-0383-z

Gettier and the method of explication: a 60 year old solution to a 50 year old problem

Erik J. Olsson

Published online: 13 September 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract I challenge a cornerstone of the Gettier debate: that a proposed analysis of the concept of knowledge is inadequate unless it entails that people don't know in Gettier cases. I do so from the perspective of Carnap's methodology of explication. It turns out that the Gettier problem per se is not a fatal problem for any account of knowledge, thus understood. It all depends on how the account fares regarding other putative counter examples and the further Carnapian desiderata of exactness, fruitfulness and simplicity. Carnap proposed his methodology more than a decade before Gettier's seminal paper appeared, making the present solution to the problem a candidate for being the least ad hoc proposal on the market, one whose independent standing cannot be questioned, among solutions that depart from the usual method of revising a theory of knowledge in the light of counterexamples. As an illustration of the method at work, I reconstruct reliabilism as an attempt to provide an explication of the (externalist) concept of knowledge.

Keywords Gettier problem · Explication · Rudolf Carnap · Reliabilism · Knowledge · Philosophical methodology

1 Introduction

There is hardly any contemporary epistemologist who has not spent some time pondering the problem posed by Edmund Gettier in his seminal article from 1963. In his Brown in Barcelona example we are invited to consider the following propositions:

E. J. Olsson (✉)
Department of Philosophy, Lund University, 221 00 Lund, Sweden
e-mail: erik_j.olsson@fil.lu.se

(i) Jones owns a Ford
(ii) Brown is in Barcelona
(iii) Jones owns a Ford or Brown is in Barcelona

Smith believes (i) with justification. He does not believe, nor is he justified in believing, (ii). Smith now deduces (iii) from (i) on purely logical grounds. Being justified in believing (i), he is also justified in believing (iii). As it happens, (i) is false but, luckily for Smith, (ii) turns out to be true, whence (iii) is still true. Hence, Smith's belief in (iii) constitutes knowledge on the respectable JTB (justified true belief) analysis. But, Gettier claims, and most commentators have concurred, given the circumstances, (iii) is not knowledge. The conventional moral drawn, following Gettier's own analysis, is that having a justified true belief that p does not imply knowing that p, and that the JTB analysis of knowledge, which states otherwise, is in error.

There is no shortage of reactions to the Gettier problem. Thus, epistemologists have suggested revising the JTB account, e.g. by adding a fourth stating that the conclusion must not rely on false lemmas (Clark 1963) or, alternatively, that the justification must be in some sense indefeasible (e.g. Lehrer and Paxson 1969). Others have proposed more radical departures from JTB, e.g. by advocating a causal or relevant alternatives account (see Goldman 1967, 1976, respectively). A third strategy amounts to the suggestion that knowledge, following Dretske (1971) and perhaps more famously Nozick (1981), should satisfy some counterfactual condition that rules out Gettier cases. One such condition is that of safety, stating that the belief in question could not easily have been false (e.g. Sosa 1999). In an influential paper, Zagzebski (1994) argued, more pessimistically, that "Gettier problems are inescapable for virtually every analysis of knowledge which at least maintains that knowledge is true belief plus something else" (p. 65). Williamson (2000) reached a similar dismal conclusion and took the practical lesson to be that epistemologists should regard the concept of knowledge as a primitive rather than attempting to analyze it in terms of other notions. This is Williamson's radical "knowledge first" approach to epistemology. I am sympathetic to Zagzebski's and Williamson's reasons for thinking that the Gettier problem is unlikely to be solved in the standard way of revising a theory in response to counterexamples, a methodology which so far has produced anything like a consensus. Hence, I will be looking for alternatives to that approach.

A common assumption in gettierology is the following:

(A) Pre-systematic judgment universally reports that people don't know in Gettier cases.

On the basis of (A), it is concluded that

(B) a proposed analysis of knowledge is inadequate unless it entails that people don't know in Gettier cases.

There has been considerable debate concerning the validity of (A), not least from the experimental philosophy camp (Weinberg et al. 2001; Swain et al. 2008). However, I will assume for the sake of the argument that (A) is true. What I will question is rather the tenability of (B). I will argue that even an analysis of knowledge that implies that people know in Gettier cases may be, in a sense, adequate and useful.

Specifically, I will question (B) on the basis of methodological considerations that were advanced a decade before Gettier's paper was published and endorsed by both Carnap (1950) and Quine (1960), i.e. by two of the founding fathers of analytic philosophy as it is currently practiced. By contrast, other non-standard approaches to the Gettier problem that I am aware of were generally formulated after, and in response to, Gettier's paper (or prior to it but by philosophers of lesser standing).[1] Whatever problems the present proposal might have, adhockery is surely not one of them. None of this means, of course, that other proposals are ad hoc in any absolute sense. I am operating here with a notion of degree of adhockery.

The methodology I favor will be introduced in Sect. 2. In Sect. 3, I consider the consequences if adopting this framework for the Gettier problem. In Sect. 4, I apply the methodology to the particular case of the reliabilist theory of knowledge. This will serve, I hope, to make the ramifications of the former more vivid.

2 Explicative epistemology

The method of explication was introduced by Rudolf Carnap, most systematically in Carnap (1950), as a procedure for defining scientific concepts in general, and philosophical concepts in particular. Quine defends Carnap's account in the closing chapter of his 1960 essay. By the procedure of explication, Carnap explains, "we mean the transformation of an inexact, prescientific concept, the *explicandum*, into a new exact concept, the *explicatum*" (p. 3). Thus an explication can be seen as a function or mapping from an informal domain to a formal, exact domain. The fact that the explicandum is not precise has the important consequence that "if a solution for a problem of explication is proposed, we cannot decide in an exact way whether it is right or wrong" (p. 4). Instead of asking whether the solution is right or wrong we should ask whether it is satisfactory, fruitful, and so on.

The method proceeds in two steps:

(1) The clarification of the explicandum
(2) The specification of the explicatum

The rationale for the first step is that a given term may have many different meanings in ordinary language. Hence unless the intended meaning is clearly identified from the start, it is unlikely that the method of explication will yield a useful result. Therefore, "[a]lthough the explicandum cannot be given in exact terms, it should be made as clear as possible by informal explanations and examples" (p. 3).

[1] There is an exception to this rule, namely, an infallibilist theory according to which knowledge requires infallible, or conclusive, justification. As epistemologists tend to agree, Gettier cases cannot arise if knowledge is assumed to require infallible justification. Moreover, infallibilism has respectable pre-Gettier pedigree in Cartesian philosophy (see Kaplan 1985, for a discussion of this point). Apart from Dretske (1971) few contemporary epistemologists have been inclined to proceed in this direction. A comparison of the relative merits of infallibilism and the response to the Gettier problem proposed here is at any rate beyond the scope of the present article.

Once the first step has been taken, there remains the task of specifying the explicatum, the exact concept which is to replace the explicandum, at least in certain contexts. Carnap sometimes uses the term "explication" to denote the definitorial rule by means of which the explicatum is specified. As we will see, Carnap states a number of desiderata on an explicatum: that it be similar to the explicandum, exact, fruitful and simple. We will consider these central conditions in a moment.

As Carnap notes, there is a tendency to think that, since the explicandum cannot, by definition, be given in exact terms anyway, it does not matter much how it is specified. But, he continues, this inference would be "quite wrong":

> On the contrary, since even in the best case we cannot reach full exactness, we must, in order to prevent the discussion of the problem from becoming entirely futile, do all we can to make at least practically clear what is meant as the explicandum. What X means by a certain term in contexts of a certain kind is at least practically clear to Y if Y is able to predict correctly X's interpretation for most of the simple, ordinary cases of the use of the term in those contexts. It seems to me that, in raising problems of analysis or explication, philosophers very frequently violate this requirement. They ask questions like: 'What is causality?', 'What is life?', 'What is mind?', 'What is justice?', etc. Then they often immediately start to look for an answer without first examining the tacit assumption that the terms of the question are at least practically clear enough to serve as a basis for an investigation, for an analysis or explication.

Hence, the fact that the terms in question are unsystematic and inexact does not mean that there is no way of approaching intersubjective agreement regarding their intended meaning. The procedure suggested by Carnap is to indicate the meaning by some well-chosen examples, perhaps supplemented by "an informal explanation in general terms" (p. 4).

To illustrate the point, Carnap invites the reader to consider the term "salt". Suppose we wish to explicate this concept. Then we might say, for example, "I mean by the explicandum 'salt', not is wide sense which it has in chemistry but its narrow sense in which it is used in the household". This explanation is an elucidation of the explicandum: a specification of what is the intended meaning of the term in question; it is not yet an explication. For this explicandum to be explicated, it is required that a suitable explicatum be found in some exact science. A possible explicatum might be "sodium chloride", or NaCl in the language of chemistry.

Or, to take Carnap's second example, suppose we wish to explicate the term "true". An elucidation of the explicandum may involve stating that we do not intend the meaning "true" has in phrases like "a true democracy" or "a true friend", but rather the meaning it has in phrases like "this sentence is true", "what he just said is true", and so on. Again, in making this statement we have not yet explicated the term "true"; we have only clarified its intended meaning. An explication would involve specifying a formal theory of truth, e.g., in logical or set-theoretical terms, perhaps drawing on the work of Tarski and others.

Clarifying the explicandum serves the purposes of specifying, perhaps in relatively crude terms, what is to be included and what is to be excluded. Once this has been accomplished we can meaningfully discuss possible explications of the term in question.

What requirements should be placed on a suitable explicatum once the explicandum has been sufficiently elucidated? The purpose of an explication is to introduce a more or less vague or unclear intuitive concept into an exact framework. Thus, we wish to find a more exact concept that in some sense corresponds to the intuitive, everyday concept. But what does this relation of correspondence entail? It is obvious that we cannot hope for complete correspondence in meaning. After all, the whole point of explication is, in a sense, to diverge from the meaning of the intuitive concept (by introducing a more exact correlate). Carnap nonetheless considers the possibility that the explicatum should be as close or similar to the explicandum as the latter's vagueness permits. However, he ultimately rejects this notion on the grounds that actual scientific practice often runs counter to it.

Take for instance the artificial scientific concept fish, defined in *Encyclopedia Britannica* as "any of more than 30,000 species of cold-blooded vertebrate animals (phylum Chordata) found in the fresh and salt waters of the world", for which Carnap uses the latin term *piscis*. This concept has come to replace the everyday concept of fish which happened notwithstanding the fact that *piscis* is a narrower concept that excludes several kinds of animal which were previously subsumed under the concept fish, e.g. whales and seals. The main reasons for the replacement is that zoologists found that *piscis* is a much more fruitful concept than fish. In general, "[a] scientific concept is the more fruitful the more it can be brought into connection with other concepts on the basis of observed facts; in other words, the more it can be used for the formulation of laws." (p. 6)

For instance, unlike whales and seals most *piscis* have a streamlined body for rapid swimming, extracts oxygen from water using gills or uses an accessory breathing organ to breathe atmospheric oxygen, have two sets of paired fins, usually one or two (rarely three) dorsal fins, an anal fin, and a tail fin, have jaws, have skin that is usually covered with scales, and lay eggs. Whales, by contrast, are taken to belong to the category of mammals because, like other mammals, they breathe air, are warm-blooded, nurse their young with milk from mammary glands, and have body hair. At the same time, *piscis*, apart from being a precise and simple concept, is sufficiently close to the common sense concept of fish to replace it, not only in scientific but also in everyday contexts.

Similar examples abound. For example, analytic geometry made it possible to introduce intuitive geometrical concepts like "line", "distance", "angle" and so on in the precise framework of arithmetic using a coordinate system. As a consequence, algebraic and analytic methods could be used to investigate these concepts and their relations in a spectacularly fruitful manner.

Generalizing from such prototypical examples of scientific definitions, Carnap arrives at the following general requirements on an explicatum (1950, p. 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 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.

Carnap adds, referring to the fourth requirement, that "[t]he simplicity of a concept may be measured, in the first place, by the simplicity of the form of its definition and, second, by the simplicity of the forms of the laws connecting it with other concepts" (1950, p. 7). We may call the first kind of simplicity *conceptual* and the second kind *theoretical* (although these terms are not used by Carnap himself in this context). He also emphasizes that simplicity is of secondary importance as many scientific theories involve complex concepts that turn out to be highly useful.

> In general, simplicity comes into consideration only in cases where there is a question of choice among several concepts which achieve about the same and seem to be equally fruitful: if these concepts show a marked difference in the degree of simplicity, the scientist will, as a rule, prefer the simplest of them. (p. 7)

As we saw, one of Carnap's examples concerns the explication of the philosophically prominent truth predicate, and the main purpose of his 1950 treatise on probability is to explicate the concept(s) of scientific confirmation central to philosophy of science. Carnap clearly thinks, then, that philosophical definitions are no different in principle from scientific ones, and that in both cases the method of explication is to be recommended. In particular, definitions of knowledge, justification and other concepts of epistemological interest are best viewed as explications. This is an attractive position for instance if one thinks of philosophy as a branch of science, or as at least closely related to science, or some variation on that theme. One prominent case in point is Quine who wrote, in the closing paragraph of *Word and Object* (Quine 1960):

> The philosopher's task differs from the others', then, in detail; but in no such drastic way as those suppose who imagine for the philosopher a vantage point outside the conceptual scheme that he takes in charge. There is no such cosmic exile. He cannot study and revise the fundamental conceptual scheme of science and common sense without having some conceptual scheme, whether the same or another one less in need of philosophical scrutiny, in which to work. He can scrutinize and improve the system from within, appealing to coherence and simplicity; but this is the theoretician's method generally. He has recourse to semantic ascent, but so has the scientist. And if the theoretical scientist in his remote way is bound to save the eventual connections with non-verbal stimulation, the philosopher in his remoter way is bound to save them too. (Quine 1960, pp. 275–276)

I will use the term *explicative philosophy* to refer to this view on the nature of philosophical definitions, in so far as they aim to define or analyze a concept occurring in natural language. I will henceforth refer to an epistemology based on this methodological position, correspondingly, as *explicative epistemology*.[2,3]

3 Implications for the Gettier problem

Suppose now that someone proposes an explication of knowledge along the following lines:

(K) S knows that p if and only if C.

Suppose, moreover, that C entails that people in fact know in Gettier cases. Does it follow that (K) fails to be a good explication of knowledge? No, and here is why: Consider the first Carnapian desideratum again:

"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."

Under what conditions would the Gettier problem be a threat to the claim that (K) 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 condition C used in (K) can be used instead. But the Gettier problem does not show this: Gettier cases are too rare to jeopardize the claim that condition C can be substituted for ordinary knowledge in most cases. Hence, the Gettier problem per se can never present a critical challenge for an explication of knowledge. An explication of knowledge can never be disregarded on the sole grounds that it fails to imply that people don't know in Gettier cases.

The reason why Gettier problems are too rare is that they involve the consecutive occurrence of two improbable events. The Brown in Barcelona example is a case in point: A proposition (Brown owns a Ford) that is strongly supported by evidence turns out, improbably, to be false and yet, improbably again (Brown is in Barcelona), the target proposition (Brown owns a Ford or is in Barcelona) comes out true anyway. Obviously the consecutive occurrence of two improbable events will be no more probable than those improbabilities individually. In the Brown case, the

[2] Keith Lehrer, in his book *Theory of Knowledge* (1990), explicitly endorses what I have called explicative epistemology as his methodological framework. However, it is doubtful whether Lehrer actually applies the methodology consistently in practice: his final definition of knowledge is anything but simple, and considerations of fruitfulness, as Carnap meant that notion to apply, do not enter prominently into Lehrer's motivations for his account.

[3] Weatherson (2003) offers a closely related account of philosophical and epistemological methodology. Just like Carnap, he gives four criteria along which a given analysis of a philosophical concept. The exact formulation of those criteria differs somewhat from Carnap's but the gist is very much the same. The view that philosophical definitions should aim at identifying natural kinds is closely related to Carnap's methodology (e.g. Kornblith 2002). See Olsson (2012) for a first sketch of the present defense of reliabilism as an explication of knowledge.

two improbable events are plausibly independent, making their consecutive occurrence even less probable. This, too, is arguably a representative feature of the example.

We must still acknowledge that if a proposed explication of knowledge gives the wrong result in Gettier cases, this can contribute to the eventual rejection of that proposal—if, that is, the analysis gives the wrong result in response to other putative counterexamples as well. Even so, the point still stands: the Gettier problem per se does not severely threaten any explication of knowledge. This is due to the fact that failure to give the right result in Gettier cases can be offset by impressive performance on other purported counterexamples or on the explicative dimensions of exactness, fruitfulness and simplicity.

4 A case study: reliabilism

The method of explication presents a general and epistemologically neutral picture from which Gettier problems can be viewed and treated. For instance, an advocate of JTB could refer to the method in a defense of her position provided she can show that JTB is exact, fruitful and simple and does well in connection with other putative counterexamples. Be that as it may, the focus of the present case study, the main purpose of which is to illustrate explicative epistemology "at work", will be on the externalist theory of (process) reliabilism.

According to 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 obtained through a reliable process. Let us suppose that the explicandum has been suitably clarified so that we have identified a pre-systematic concept of knowledge for which the reliabilists accounts seem appropriate. Presumably, this would be the (externalist) concept of knowledge as it applies not only to adult humans but also to higher level animals and children. The next step is that we satisfy ourselves that the requirements on an explicatum are satisfied to a high degree. In the case of the reliabilist account of knowledge this means the following:

(1) Reliabilist knowledge must be similar to the common concept of knowledge in such a way that, in most cases in which the latter 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, but this requirement is less important than the other three.

Following the discussion above, I assume for the sake of the argument that reliabilism gives the wrong result in Gettier cases. In other words, I assume that Gettier cases are cases in which the true belief in focus was produced by a reliable process. The validity of this assumption can be questioned. (Wasn't the inferential process leading up to Smith's belief that Jones owns a Ford or Brown is in Barcelona dependent on a false premise, and aren't inferential processes that so depend unreliable?) But in the interest of simplicity we won't question it here.

If reliabilism gives the wrong result in Gettier cases, that counts against reliabilism vis-à-vis requirement. (1) But, as we saw, this setback could be offset by impressive performance regarding (2)–(4). Let us therefore ask how well reliabilism fares along the dimensions of fruitfulness, simplicity and exactness, in that very order.

Fruitfulness is clearly a central requirement in the Carnapian explicatory framework, so this is where we start. Reliabilist knowledge, if it is to be fruitful, must be useful for the formulation of many universal statements (empirical laws, since it is arguably a non-logical concept). As I have argued at length elsewhere, reliabilist knowledge is positively correlated with:

(i) having more true beliefs (and reliabilist knowledge) in the future (Goldman and Olsson 2009),
(ii) having a more stable true belief (Olsson 2007), and
(iii) having greater chance of acting with practical success over time (Olsson 2007).

These correlations give rise to statistical laws. In the case of (ii), for instance, the corresponding statistical law could be formulated as follows:

P(S's true belief that p will stay put | S's true belief that p was obtained by a reliable process) >
P(S's true belief that p will stay put | S's true belief that p was obtained by an unreliable process)

There are, then, at least three empirical laws that involve reliabilist knowledge, namely those just mentioned concerning repetition, stability and practical success. Rather than dwelling on these matters here I refer the interested reader to the articles in which they are detailed. I conclude that reliabilist knowledge seems to satisfy Carnap's desideratum of fruitfulness to a significant degree.[4]

These statistical laws are of interest also from the point of view of the requirement of simplicity which, as Carnap pointed out, has theoretical simplicity as a part, i.e. simplicity of the laws in which the explicatum figures prominently. Clearly, the statistical laws cited above are strikingly simple, making reliabilist knowledge a simple explicatum in this sense. Simplicity, as we saw, also has a

[4] The statistical correlations referred to here were originally appealed to in the context of the so-called swamping problem for reliabilism: the problem of how to account for the greater value of items of reliabilist knowledge over the corresponding items of mere true belief (for an overview, see Olsson 2011). In the present context, however, they are used to support the value of the *concept* of reliabilist knowledge. For this reason, the present investigation is related to that of Craig (1990) which also focuses on the value of the concept of knowledge.

purely conceptual aspect having to do with the simplicity of the clauses that define the explicatum. Disregarding the conditions of truth and belief, which are common to most prominent definitions of knowledge and therefore irrelevant in a comparative context, we must ask whether the distinctive third reliability clause is simple. That clause states, we recall, that the true belief in question has been produced by a reliable process. Compared with many other accounts of knowledge, some of which takes several pages to spell out (e.g. Shope 2002, or Lehrer 1990), this clause does look attractively simple.

But first impression may mislead. For isn't the simplicity of the reliabilist account challenged by the notorious generality problem? Reliabilism does not include a method for deciding how a belief formation process should be classified and for deciding, based on that method, whether the process was reliable, making it allegedly "seriously incomplete" (Conee and Feldman 1998). Moreover: supplementing reliabilism accordingly is bound to make that theory much more complicated.

In response to this type of concern, many authors have remarked that the issue raised is not one that is unique to reliabilism (e.g. Comesana 2006; Goldman 2008; Bishop 2010). A standard JTB account, for instance, does not per se tell us whether a given belief was justified in a concrete case. Similarly, evidentialism—the view that requires of knowledge that the belief in question be based on adequate evidence—does not tell us whether a given belief was based on adequate evidence in a concrete case (Feldman and Conee 1985). While this is certainly a point worth making, noticing that other theories have a problem contributes little to solving it. I take the deeper observation to be that theories of knowledge state necessary and sufficient conditions for the applicability of a concept; they don't, nor do they have to, include additional criteria for deciding, in concrete cases, whether those conditions are satisfied. This goes in particular for theories of knowledge that are advanced as explications. Hence, the fact that reliabilism standardly does not include a method for deciding how a belief formation process should be classified is not a sign of theoretical or explicative incompleteness. Hence, there is no reason to think that additional clauses are needed that would make the pleasantly straightforward reliability condition more complex. We may conclude that the reliabilist explication of knowledge scores highly not only on fruitfulness but also on both kinds of simplicity.

Moving on to 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. Let us, once more, focus on the characteristic third clause of the reliabilist explication: "S's belief that p was acquired through a reliable process". The crucial concept here, obviously, is that of a reliable process of belief formation. Surely, this concept can be made precise within statistical theory. A process of belief formation can be modeled in statistics as a random variable: a trial is performed, typically an observation, leading the formation of a belief. This is analogous to how throwing a die gives rise to a certain number coming up. In other words, a process of belief formation can be modeled in statistics as a function from a set of trials or environmental stimuli X to a set of potential beliefs B. The reliability of such a function, with regard to a certain type of belief, can be spelled out in terms of how often, relatively speaking, the process results in beliefs of that type. In our case, the reliability of such a function with

regard to truth is determined by the frequency of trials that result in beliefs that are true.[5]

The bottom line is that the reliabilist explication of knowledge introduces the concept of knowledge into the well-connected system of scientific concepts represented by statistical (probability) theory. As a consequence, the exactness of reliabilism has been reduced to the exactness of statistical theory. If you believe that the latter is sufficiently exact, as arguably most scientists and perhaps even philosophers do, then you will have to accept that the former is sufficiently exact as well.[6]

Critics may complain that modeling a belief formation process in statistical terms presupposes that we know what kind of process we are talking about. Otherwise, we cannot identify the sample space, that is to say, we cannot say what the possible outcomes of the process are. However, this is surely a general problem in the application of statistics to real life. Take, for example, the process of casting a die. The natural sample space would be {1, 2, 3, 4, 5, 6}, corresponding to the different sides of the die that can come up, but we can also think of the process in terms of the sample space {even number, odd number} corresponding to whether an even or odd number of dots come up. If the issue of selecting a sample space negatively affects the exactness of reliabilism, then that very same issue does the same to statistics generally. But by the same token, the exactness of formal logic would be challenged by the fact that it is sometimes difficult to apply formal logic to real life arguments. At any rate, even if all we can establish is that the exactness of reliabilism has been reduced to the exactness of statistics this still seems to put reliabilism in a rather favorable light, comparatively speaking.

I have argued so far that the assumed fact that reliabilism does not give the right result in Gettier cases can be offset by the observation that the reliabilist concept of knowledge is fruitful, simple and exact (or as exact as statistical theory). If the Gettier problem were the only challenge from the perspective of similarity to the ordinary concept, we would be done. But, alas, it is not. The problems of generality, in the form it has been raised by Conee and Feldman (1998), evil demon (Cohen 1984), clairvoyance (Bonjour 1980), Truetemp (Lehrer 1990), and bootstrapping (Cohen 2002) all fit into this category.[7] Space does not permit an investigation into all these alleged difficulties. In the following, I will focus on Conee and Feldman's version of the generality problem viewed as challenging the similarity of reliabilist knowledge to the ordinary concept.

What needs to be shown is 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 Conee and Feldman (1998) have argued, as I understand

[5] This sketch of a proposal for how to model the concept of a belief forming process within statistical theory draws on the lucid discussion in Levi (2003, pp. 144–150). Levi argues that a belief forming method can be thought of as a method of routine decision making.

[6] The point made here is related to Adler and Levin's remark that the generality problem is "too general" (Adler and Levin 2002).

[7] Some of these problems were raised for reliabilism as a theory of justification rather than as a theory of knowledge.

them.[8] Conee and Feldman's argument, put forward in the context of a discussion of Heller's (1995) claim that people typically manage to converge on the same process categorization and reliability judgment, proceeds in the following steps:

CF1. People will agree on process typing and reliability only if a particular process type is salient in the conversational context.
CF2. A particular process type is salient in the conversational context only if it has been explicitly mentioned.
CF3. In many cases in which no type has been explicitly mentioned, we still agree on the corresponding knowledge claims.
CF4. Hence, in many cases people will agree on the corresponding knowledge claims without agreeing on process typing and reliability.
CF5. 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 from CF5 that the reliabilist concept of knowledge is not sufficiently similar to the ordinary concept of knowledge and that it fails to satisfy Carnap's first criterion on an explication of knowledge.

However, as pointed out in Olsson (in press), premise CF2, the claim that a particular process is salient in a conversational context only if it has been explicitly mentioned, is incongruent with the findings of the influential basic level tradition in cognitive psychology (e.g. Rosch et al. 1976). A cornerstone of that tradition is that the level of categorization is determined by two cognitive forces. On the one hand, we strive to maximize economy. This desideratum favors a few broad or general categories. On the other hand, we wish to maximize informativeness, i.e., we want to be able to infer many useful things from the fact that something belongs to a given category. This desideratum favors narrow or specific categories. The optimal category turns out to be mid-range categories representing a reasonable compromise between the two goals. Thus, I tend to classify the thing I am sitting on right now as "chair" rather than broadly as "furniture" or narrowly as "kitchen chair".

Until recently, however, the basic level theory had apparently not been directly tested for belief formation categorization. As a remedy, Jönsson (2013) asked participants to respond to film clips of a person's acquiring a belief by categorizing the process. Importantly, no process type was explicitly mentioned in the clips or in the instructions. Even so, Jönsson found there to be a strong correlation between how subjects categorized the processes. The categories subjects converged on were midrange categories such as "seeing", "hearing", and so on. Notwithstanding Conee and Feldman's prediction, these categories were apparently salient for the subjects in spite of the fact that they were not explicitly mentioned.

It is worth noting that this sketch of a defense of reliabilism as an explication of knowledge in the face of its assumed failure to deal with the Gettier problem relied entirely on Carnap's own formulation of the requirements on an explication. At no

[8] See also Feldman and Conee (2002). Given the assumption of the present article that reliabilism does not give the right result in Gettier cases, a weaker conclusion would suffice for Conee and Feldman's purposes. However, this fact does not affect the present line of argument.

point did I depart, modify or in other ways fiddle around with these requirements. In particular, fruitfulness was understood, following Carnap, as being strictly a matter of the explicatum figuring prominently in empirical laws (since we are dealing with an empirical concept). The defense was, in this sense, splendidly non-adhoc.[9]

5 Discussion and conclusion

Despite its many merits, Carnap's method of explication has not been without its critics. In his contribution to the Schilpp (1963) volume on Carnap's philosophy, Strawson objected that offering explications for those who seek philosophical illumination is irrelevant and changes the subject. In his own words:

> It seems prima facie evident that to offer formal explanations of key terms of scientific theories to one who seeks philosophical illumination of essential concepts of non-scientific discourse, is to do something utterly irrelevant—is a sheer misunderstanding, like offering a textbook on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart. (1963, p. 505)

In his reply to Strawson, Carnap drew on the following analogy:

> A natural language is like a crude, primitive pocketknife, very useful for a hundred different purposes. But for certain specific purposes, special tools are more efficient, e.g., chisels, cutting machines, and finally the microtome. If we find that the pocket knife is too crude for a given purpose and creates defective products, we shall try to discover the cause of the failure, and then either use the knife more skillfully, or replace it for this special purpose by a more suitable tool, or even invent a new one. [Strawson's] thesis is like saying that by using a special tool we evade the problem of the correct use of the cruder tool. But would anyone criticize the bacteriologist for using a microtome, and assert that he is evading the problem of correctly using the pocketknife? (1963, p. 938f.)

This reply to Strawson assumes that important philosophical problems remain unsolved because natural language is too crude for addressing them. For the purposes of those problems, we should use more precise (logical or mathematical) tools. The philosopher who fails to see this is like the bacteriologist who uses a pocketknife rather than a microtome to solve a specialized problem of bacteriology. However, as Maher notices, "the relevant analogy for 'one who seeks philosophical illumination of essential concepts of non-scientific discourse' is someone who seeks knowledge of proper use of the pocket knife" (Maher 2007, p. 333). If so, "Carnap has offered nothing to satisfy such a person" (ibid.).

[9] In a defense of JTB along similar lines an enriched conception of fruitfulness, including ethical or legal fruitfulness, would probably be useful. This would involve a slight departure from Carnap's original formulation of his methodology. Weatherson (2003) focuses on JTB in his related methodological study.

Maher still believes that the methodology of explication can be defended against Strawson's critique. His suggestion is illuminating for several reasons, not least because it illustrates the important role of the first requirement on an explicatum, that it be similar to the explicandum. Here is what Maher proposes:

> Suppose our problem is to determine whether or not some sentence *S* of ordinary language is true. If we apply the method of explication to this problem, we will construct explicata for the concepts in *S*, formulate a corresponding sentence *S′* using these explicata, and determine whether or not *S′* is true. This does not *by itself* solve the original problem—that is Strawson's point—but it can greatly assist in solving the problem, in three ways. (1) The attempt to formulate *S′* often shows that the original sentence *S* was ambiguous or incomplete and needs to be stated more carefully. (2) If the explicata appearing in *S′* are known to correspond well to their explicanda in other cases, that is a reason to think that they will correspond well in this case too, and hence to think that the truth value of *S* will be the same as that of *S′*. (3)We can translate the proof or disproof of *S'* into a parallel argument about the corresponding explicanda and see if this seems to be sound; if so, we obtain a direct argument for or against *S*. In these ways, explication can provide insights and lines of argument that we may not discover if we reason only in terms of the vague explicanda. (pp. 333–334)

As an illustration, Maher takes Nicod's claim that a law of the form "All F are G" is made more probable, or confirmed, by evidence that something is both F and G. The problem is to determine whether this is true, which is a problem stated in natural language. Following the methodology of explication, the problem is solved by (1) explicating the concept of scientific confirmation, (2) showing that Nicod's claim does not hold for the explicatum and (3) arguing that this result plausibly carries over to the original problem via the similarity between the explicandum and explicatum. Of course, the more similar the explicatum is to the explicandum, the greater is the force this argument is likely to have.

Summing up, I challenged the claim that a proposed analysis of knowledge is inadequate unless it entails that people don't know in Gettier cases. I did so from the perspective of Carnap's method of explication which was originally advanced, a decade prior to Gettier's influential paper, as a general method for philosophical analysis. Other popular approaches that I know of were devised after, and in response to, Gettier's article, making them more vulnerable to criticism for adhockery. From the explicationist perspective, the Gettier problem per se is not a fatal problem for any explication of knowledge. It all depends on how the explication fares regarding other putative counter examples, and—importantly—the further desiderata of exactness, fruitfulness and simplicity. The main insight is that we should care, perhaps more than we normally do, about whether an analysis of knowledge gives rise to a concept that is fruitful in the sense that it figures prominently in law-like statements. The present approach has the further advantage of being (essentially) epistemically neutral e.g. vis-à-vis the internalist–externalist distinction, making it available to epistemologists with widely different outlooks. In an attempt to add vividness to the exposition and illustrate the method of explication

at work, I devoted some space to the assessment of reliabilism from the point of view of explicative epistemology. My thesis in this connection was that even if the reliabilist explication of knowledge fails regarding Gettier, it is a promising explication with respect to Carnap's desiderata, and it can also handle the generality problem in its various guises, including the ingenious—but as it turns out empirically unsound—argument put forward by Conee and Feldman (1998).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Adler, J., & Levin, M. (2002). Is the generality problem too general? *Philosophy and Phenomenological Research, 65*(1), 87–97.

Bishop, M. A. (2010). Why the generality problem is everybody's problem. *Philosophical Studies, 151*, 285–298.

BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy, 5*, 53–73.

Carnap, R. (1950). *Logical foundations of probability*. Chicago: Chicago University Press.

Carnap, R. (1963). Replies and systematic expositions. In P. A. Schilpp (Ed.), *The philosophy of Rudolf Carnap* (pp. 859–1013). La Salle, IL: Open Court.

Clark, M. (1963). Knowledge and grounds: A comment on Mr. Gettier's Paper. *Analysis, 24*, 46–48.

Cohen, S. (1984). Justification and truth. *Philosophical Studies, 46*, 279–295.

Cohen, S. (2002). Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research, 65*, 309–329.

Comesana, J. (2006). A well-founded solution to the generality problem. *Philosophical Studies, 129*, 27–47.

Conee, E., & Feldman, R. (1998). The generality problem for reliabilism. *Philosophical Studies, 89*, 1–29.

Craig, E. (1990). *Knowledge and the state of nature: An essay in conceptual synthesis*. Oxford: Clarendon Press.

Dretske, F. (1971). Conclusive reasons. *Australasian Journal of Philosophy, 49*, 1–22.

Feldman, R., & Conee, E. (1985). Evidentialism. *Philosophical Studies, 48*, 15–34.

Feldman, R., & Conee, E. (2002). Typing problems. *Philosophy and Phenomenological Research, 65*(1), 98–105.

Goldman, A. I. (1967). A causal theory of knowing. *Journal of Philosophy, 64*, 357–372.

Goldman, A. I. (1976). Discrimination and perceptual knowledge. *Journal of Philosophy, 73*, 771–791.

Goldman, A. I. (2008). Reliabilism. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Summer 2012 ed.). http://plato.stanford.edu/entries/value-intrinsic-extrinsic/.

Goldman, A. I., & Olsson, E. J. (2009). Reliabilism and the value of knowledge. In A. Haddock, A. Millar, & D. H. Pritchard (Eds.), *Epistemic value* (pp. 19–41). Oxford: Oxford University Press.

Heller, M. (1995). The simple solution to the generality problem. *Noûs, 29*, 501–515.

Jönsson, M. L. (2013). A reliabilism built on cognitive convergence: An empirically grounded solution to the generality problem. *Episteme, 10*, 241–268.

Kaplan, M. (1985). It's not what you know that counts. *Journal of Philosophy, 82*, 350–363.

Kornblith, H. (2002). *Knowledge and its place in nature*. Oxford: Oxford University Press.

Lehrer, K. (1990). *Theory of knowledge*. Boulder, CO: Westview.

Lehrer, K., & Paxson, T. D. (1969). Knowledge: Undefeated justified true belief. *Journal of Philosophy, 66*, 225–237.

Levi, I. (2003). Contracting from epistemic hell is routine. *Synthese, 135*, 141–164.

Maher, P. (2007). Explication defended. *Studia Logica, 86*, 331–341.

Nozick, R. (1981). *Philosophical explanations*. Cambridge, MA: Harvard University Press.

Olsson, E. J. (2007). Reliabilism, stability, and the value of knowledge. *American Philosophical Quarterly, 44*, 343–355.

Olsson, E. J. (2011). The value of knowledge. *Philosophy Compass, 6*, 874–883.

Olsson, E. J. (2012). Reliabilism as explicating knowledge: A sketch of an account. In C. Jäger & W. Löffler (Eds.), *Epistemology: Contexts, Values, Disagreement*. Germany: Ontos Verlag.

Olsson, E. J. (in press). A naturalistic approach to the generality problem. In H. Kornblith & B. McLaughlin (Eds.), *Goldman and his critics*. Hoboken, NJ: Wiley.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.

Schilpp, P. A. (1963). *The philosophy of Rudolf Carnap*. LaSalle, IL: Open Court.

Shope, R. K. (2002). Conditions and analyses of knowing. In P. K. Moser (Ed.), *The Oxford handbook of epistemology* (pp. 25–70). Oxford: Oxford University Press.

Sosa, E. (1999). How to defeat opposition to Moore. *Philosophical Perspectives, 13*, 141–154.

Strawson, P. F. (1963). Carnap's views on constructed systems versus natural languages in analytic philosophy. In P. A. Schilpp (Ed.), *The philosophy of Rudolf Carnap* (pp. 503–518). La Salle IL: Open Court.

Swain, S., Alexander, J., & Weinberg, J. (2008). The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research, 126*(1), 138–155.

Weatherson, B. (2003). What good are counterexamples? *Philosophical Studies, 115*, 1–31.

Weinberg, J. M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29*, 429–460.

Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.

Zagzebski, L. (1994). The inescapability of getter problems. *The Philosophical Quarterly, 44*(174), 65–73.