Conceptual Construction in Epistemology:
Why the Content of Our Folk Terms Has Only Limited Significance

Thomas Grundmann
University of Cologne

Abstract. Standard Analytic Epistemology typically relies on conceptual analysis of folk epistemic terms such as 'knowledge' or 'justification'. A cross-cultural and cross-linguistic perspective on this method leads to the worry that there might not be universally shared epistemic concepts, and that different languages might use folk notions that have different extensions. Moreover, there is no reason to believe that our epistemic common-sense terms pick out what is epistemically most significant or valuable. In my paper, I take these issues as a starting point for exploring the prospects of an alternative methodological approach that I call 'alethic instrumentalism'. The core idea behind this approach is to start with a properly designed epistemic goal and then to develop a framework of instrumentally valuable methods oriented towards this goal. This results in a somewhat revisionary framework of newly constructed core epistemic terms. In the paper, I elucidate the foundations of this new framework and address a number of methodological and content-related objections to the approach.

Standard Analytic Epistemology (SAE) has dominated epistemology in the English-speaking world for much of the 20th century and beyond (Bishop and Trout 2005, p. 8). SAE can be characterized by its use of a specific method, namely the method of cases. Typically, its proponents substantiate or reject general accounts of knowledge or (epistemic) justification by testing whether they accord with their intuitive attributions of 'knowledge' or 'justification' to particular cases.[1] This method treats intuitions about particular cases as evidence—evidence that is based on the individual cognizer' competence with respect to her use of the relevant epistemic concepts. The thought is that she can use her conceptual competence to elucidate what knowledge and justification in fact are. Accordingly, conceptual analysis provides us with a means to articulate epistemic reality in itself, and not just our representations of it.[2]

Currently, dominant views in philosophy of language and philosophy of mind cast some doubt on the reliability of conceptual analysis. For example, externalist views of content challenge the idea that the cognizer's application-dispositions properly reflect the concept's actual application-conditions (Putnam 1975). Other concerns relate to SAE more directly. In this respect, three challenges are of major interest here: (a) the challenge from cross-cultural variation of epistemic intuitions, (b) the challenge from cross-linguistic diversity of epistemic terms, and (c) the challenge from the insignificance of epistemic folk concepts.

In my paper, I take these challenges as a starting-point for exploring the prospects of an alternative methodological approach that I call 'alethic instrumentalism'.[3] The core idea behind the approach defended here is to start with a properly designed epistemic goal and then to develop a framework of instrumentally valuable methods oriented towards this goal. This results in a somewhat revisionary framework of newly constructed core epistemic terms. In the first section, I examine how serious the three above-mentioned challenges are. Taken together, they sufficiently motivate the search for alternatives to SAE. In the second section, I discuss some standard alternatives to SAE and consider their problems and limitations. In the

---

[1] See Grundmann and Horvath 2014 and Grundmann 2017 for a more comprehensive characterization of this method.

[2] See Goldman 2007 for a dissenting view.

[3] Somewhat similar approaches are suggested by Quine 1986 (pp. 664–5) and by Bishop and Trout 2005.

third section I elucidate the foundations of alethic instrumentalism, and in the fourth section, I develop it in more detail. I proceed by addressing a number of methodological and content-related objections to this approach, and conclude with some general considerations.

1. Three Challenges to SAE

Proponents of SAE rely on the methodological assumption that their individual intuitions about epistemically relevant cases can be generalized. In their seminal 2001 paper, Weinberg, Nichols, and Stich presented evidence that suggests that intuitions about Gettier-cases are not sufficiently stable to ground epistemology. The competent speakers of English that participated in this study varied in their judgments about Gettier-like cases depending on their cultural background. Whereas the majority of undergraduates from Rutgers University with a Western background judged that knowledge is absent in the Gettier case and hence reproduced the textbook view, the majority of competent speakers with an East Asian background claimed that knowledge prevails in the Gettier case (Weinberg, Nichols, and Stich 2001). Provided one allows that the extension of the English verb 'know' does not differ when uttered by competent speakers with different cultural backgrounds, the study suggests that epistemic intuitions are strongly influenced by factors that are irrelevant to the truth of knowledge-attribution. This would threaten the trustworthiness of epistemic intuitions.

How serious is this **challenge from cross-cultural variation of epistemic intuitions**? Interestingly, in more recent studies the cross-cultural variation of intuitions about Gettier-cases could not be replicated (see Nagel, Mar, and San Juan 2013, Seyedsayamdost 2015, Kim and Yuan 2015). Although this sounds like good news for proponents of SAE, it seems too early to come to a final verdict on this issue. Intuitions in general seem to be influenced by a variety of irrelevant factors (Alexander 2012). Whether cultural background is among them, and whether this influence is strong enough to undermine the trustworthiness of intuitions even under favourable conditions, is still an open question. So it seems fair to conclude that at least a lingering worry remains from the initial challenge.

While the challenge from cross-cultural variation of epistemic intuitions concerns intuitional variation among competent speakers of the *same language*, typically English, the **challenge from cross-linguistic diversity of epistemic terms** concerns extensional differences between related locutions in *different languages*. Here is a recent articulation of this worry by Stich and Mizumoto (forthcoming, pp. 3–4):

> Contemporary English is one of approximately 6,000 languages spoken in the world. It is the native language of less than 6% of the world's population. Moreover, when Western epistemology emerged, in Ancient Greece, English did not exist. So why should the usage of sentences in contemporary English in which 'know' and cognate words occur, the concept of knowledge expressed by those words, or features of the language of knowledge attribution in contemporary English, play any special role in epistemology? […] Is there any justification for this practice? One possible answer invokes what might be called the 'Universality Thesis', which claims that the properties of the English word 'know', English sentences of the form 'S knows that p' and related locutions that have been studied by Anglophone epistemologists are shared by the standard translations of these expressions in most or all languages. If this were true, then the

focus on English would simply be a matter of convenience for the vast majority of analytic epistemologists who are native speakers of, or fluent in English.

To see what is at issue here, suppose that the Universality Thesis is false. In this case, different languages will contain epistemic terms that never share their meanings and extensions completely. There wouldn't exist any Chinese, Swahili, or Hindi term that has the same extension as the English word 'knowledge'. If this were true, we would face the following dilemma: we can either use the standard for the English term 'knowledge' as the universally legitimate epistemic standard, or we can relativize epistemic standards to languages. But choosing the first horn would be completely arbitrary. Why should the standard related to the English term rather than to the epistemic terms of any other language have universal significance? (Stich 1998, p. 107). However, if we choose the second horn, epistemology becomes completely parochial (Bishop and Trout 2005, p. 108). This would result in a plurality of language-specific ethno-epistemologies. The underlying idea has attracted epistemologists who are sensitive to cultural diversity (see Mizumoto forthcoming). But why should one care about what standards apply to knowledge in one's own or someone else's language? Prima facie, these standards do not have any more normative force than conventions that are accepted within a specific social group. It seems completely contingent that we in fact follow the rules of one particular group rather than any other. Hence, understanding epistemology along the lines of ethno-epistemology would undermine the normatively binding force of epistemology in general.

The prospects of SAE therefore depend on the correctness of the Universality Thesis. If every language contains epistemic vocabulary with similar semantic properties, it is likely that epistemic terms express universal epistemic concepts. Then we can use our native language to analyze these concepts. But is the Universality Thesis correct? The body of experimental cross-linguistic studies of epistemic locutions is rapidly growing, and I cannot offer a representative survey in this paper. Rather, I will report some data from three recent studies. First, Machery et al. (2015) have tested speakers of Japanese, Bengali, Portuguese, and English on their Gettier intuitions. Surprisingly, locutions for knowledge in all four languages exhibit an anti-luck condition. This suggests, as Machery et al. (2015, p. 8) put it, 'that Gettier intuitions may be a reflection of an underlying innate and universal core folk epistemology'. Second, Mizumoto (forthcoming) conducts a cross-linguistic study of locutions for knowledge in Japanese and English. Accordingly, Japanese has two knowledge verbs rather than one: 'shitte-iru' and 'wakatte-iru'. The former expresses receiving information, the latter refers to a deeper understanding that is agent-based. A cross-linguistic comparison between responses of Japanese and English speakers to philosophically crucial cases (like Gettier, TrueTemp, fake barn, bank case etc.) shows that 'wakatte-iru' but not 'shitte-iru' displays the same features as the English verb 'know' (Mizumoto forthcoming, p. 35). Does this suggest that 'wakatte-iru' is the Japanese counterpart to the English verb 'know' and that the Universality Thesis is correct? Mizumoto questions this claim. According to him, 'know' in English might equally well be ambiguous between 'shitte-iru' and 'wakatte-iru', such that different English speakers rely on different meanings of this term when they form apparently conflicting judgments about the crucial cases (p. 36). However, I do not see why the data call for this explanation, since even users of the Japanese 'wakatte-iru' exhibit somewhat conflicting judgments about the relevant cases. So, we need not invoke the ambiguity-thesis to explain unstable linguistic

behaviour. As far as I can see, the data support the Universality Thesis, at least to some degree. Third, Matthewson and Glougie (forthcoming) have discovered that the linguistic tracking mechanisms, as well as the role and the standards of justification for assertion, are similar across native American languages and English. This study provides further evidence in support of the Universality Thesis, although a cross-linguistic study of the semantic properties of justificatory terms is still pending.

The collected data provide evidence for the truth of the Universality Thesis with respect to epistemic terms. However, at this stage the body of evidence is too small to justify the claim that this thesis has been sufficiently confirmed. The studies conducted so far need to be replicated, and cross-cultural studies of further languages and of further semantic aspects of the relevant vocabulary are required. The challenge from cross-cultural diversity of epistemic terms has not yet been fully answered.

Let us now address the third threat to SAE: **the challenge from the insignificance of epistemic folk concepts**. Here is one way of setting up this challenge: suppose that there is basically no cultural variance of intuitions about knowledge or justification and that epistemic terms for knowledge and justification have similar semantic properties across most or all languages. One might still wonder whether and why epistemology should be built upon universally shared folk terms. The worry is that these terms might not refer to what is epistemically most relevant and valuable. Suppose that linguistic analysis finally establishes that the English term 'knowledge' (and its counterparts in other languages) refers to true belief (see Sartwell 1992; Goldman 1999, pp. 23–25, calls this 'weak knowledge') or is even non-factive (as suggested by Hazlett 2010). Then it would make little sense to assess epistemic achievements only in terms of knowledge. Better standards seem available, even if they are not expressed by any of the epistemic core folk terms. Or, to choose another example, suppose that the English term 'justification' (or its counterparts in other languages) refers to whatever is licensed by the epistemic standards that are accepted within the cognizer's social environment (Goldman 1988 calls this 'weak justification'). You still might want to say that we could and should criticize the standard of the group from an epistemic point of view. Again we would take a stance that could not be articulated in terms of epistemic core folk vocabulary.

Epistemologists should not worry too much that the folk concepts may not carve at nature's joints. This worry would presuppose that epistemic affairs are completely objective and unconnected to the goals and concerns of the cognizer. Whereas such a view may be plausible for physics or metaphysics, it seems inapt for epistemology. But is it not then perfectly reasonable to assume that the universality of epistemic folk terms indicates that they are suitable to pick out those things that epistemically really matter to us? Not so fast. The universality of folk epistemic concepts makes it likely that these concepts pick out things that are anthropologically important. However, it does not suggest that the things picked out by these terms are of major *epistemic* importance. For example, our epistemic folk terms may not be *purely* epistemic. Theories of practical encroachment (Stanley 2005; Fantl and McGrath 2009, pp. 194–212) suggest that the attribution of knowledge (or justification) partly depends on how important the truth is for the believer. Impure categorizations like these may be important in practice, but they do not necessarily indicate *epistemological* importance. Other non-epistemic factors may also be relevant to our folk categorization. It is often highly

desirable to possess reasons for one's assertions that can be articulated in conversation. But it is less clear whether this internalist requirement is an epistemic requirement rather than a purely dialectical desideratum. Folk epistemic concepts may also be the result of compromising strategies that combine unrelated and sometimes even conflicting conditions in a single concept. For this reason, epistemic folk concepts may be hybrid or even incoherent. For example, according to Beckermann (2001) our concept 'knowledge' merges the epistemic goal of truth with an appropriate epistemic means towards it.[4] You might also think that our concept 'justification' combines internalist standards of first-person guidance with externalist standards of third-person assessment. Finally, one can take a genealogical stance on our epistemic folk concepts (see Craig 1990). One then raises questions such as the following: What was the point of introducing the concept—of knowledge, for instance—in the first place? Most likely, it was meant to flag certain epistemically desirable states or valuable sources of information that are achievable and often achieved in real life. If this is the correct explanation of why we have the concept of knowledge then one should expect that only minimal standards are related to it. Ideal or optimal standards would conflict with the achievability of knowledge. And there is no point in introducing a concept that does not refer to anything that is actually instantiated. However, if 'knowledge' only involves minimal standards, then you might have theoretical reasons to also introduce higher standards that are not expressed by the existing epistemic folk terms. Suppose that someone told Descartes that our ordinary term 'knowledge' refers to true belief without requiring certainty. Surely, he would not give up his high standards of assessment. Most probably, he would reply that our ordinary term is of limited significance and should be replaced by another term that properly reflects the relevant standards of epistemic evaluation.[5] The preceding criticism of folk epistemology relies on an epistemic point of view that is not shaped by our folk terms. How is that point of view accessible to us? We need to have a grasp of our guiding epistemic values and goals that is independent of currently used folk concepts.[6]

What is the upshot of the above discussion of the three major challenges to SAE? Although there is some evidence that supports the view that epistemic intuitions (about knowledge and justification) do not vary with cultural background, and although there is further evidence that epistemic terms from different languages categorize epistemic reality in roughly the same way, it is still an open empirical question whether epistemic intuitions and epistemic terms are cross-culturally sufficiently robust to provide us with a stable and solid foundation for epistemology. Moreover, epistemic folk concepts need not pick out the epistemically most relevant and most valuable properties. Taken together, these concerns about SAE motivate the search for a new approach to epistemology that is methodologically independent of our folk linguistic and conceptual resources.

2. Typical Alternatives to SAE

---

[4] Fassio and McKenna 2015 argue that the folk concept of knowledge is inconsistent because it is governed by two rules one of which is sensitive whereas the other is insensitive to stakes (p. 762).

[5] Pasnau 2013, pp. 1000–1, argues that Descartes endorses a theory of epistemic ideals that are unconnected to the extension of the English folk term of knowledge.

[6] Thanks to Joachim Horvath, Jens Kipper and Chris Ranalli for pressing me on this point.

One approach to epistemic reality that does not rely on the analysis of our epistemic folk concepts is **full-blown naturalism** (see Kornblith 2002). This kind of naturalism combines metaphysical and methodological strands. According to the metaphysical aspect of full-blown naturalism, knowledge constitutes a natural kind with a certain explanatory power. In particular, knowledge plays a relevant role in behavioural explanations of cognitive ethology. According to the methodological aspect of full-blown naturalism, our concept of knowledge is a natural kind term that is neither very informative nor trustworthy with respect to the nature of its referents. Proponents therefore first select paradigm cases that clearly belong within the extension of this term, and then examine the underlying nature of these cases empirically. Obviously, this approach bypasses any reliance on epistemic intuitions and conceptual competence and is therefore a genuine alternative to SAE.

I have two main concerns about this naturalistic approach to knowledge.[7] Both of them mirror the above challenges to SAE. First, there is the worry that what counts as paradigmatic (for knowledge) may be controversial between different cultures. It might even be controversial within our own Western culture (see Alston 2005, p. 25). Consider, for instance, sceptics and Mooreans: whereas the sceptic will claim that there is no paradigmatic case of knowledge, Mooreans will maintain that there are all kinds of cases of mundane knowledge, such as of course that one has hands.[8] The same seems to be true for internalists and externalists: whereas the former will take cases of reflective knowledge as paradigmatic, the latter will claim that what Sosa calls 'animal knowledge' is paradigmatic for knowledge. It seems likely that controversies about paradigmatic cases across cultures will be even more serious and abundant. If this is true, the choice of epistemic paradigm cases is too unstable to rest epistemology on it. Second, suppose that people broadly agree about the cases that are paradigmatic for knowledge, and that knowledge constitutes a natural kind that causally explains successful (animal) behaviour. One might still ask why this should matter for our *epistemic* assessment of knowledge. That knowledge is something epistemic that plays a robust causal role does not by itself show that knowledge is among the epistemically most important and valuable things. It rather shows that knowledge is important in other than epistemic respects, namely as a causal factor in the natural world. So, the challenge from insignificance recurs.

Can we make progress by aiming for **Carnapian explications** instead? Carnap (1950, Ch. 1) introduced this idea to recommend a specific procedure for concept transformation: we should change our pre-theoretic and inexact concepts to adjust these concepts to scientific standards and make them properly usable within scientific contexts. This transformation was originally understood as an intentional regulation of a given concept's use. Since this might be hard to achieve—given externalist views of content—one might reinterpret the basic idea such that it recommends a replacement of a pre-theoretic concept by a technical term within

---

[7] Further objections can be found in Horvath 2016, namely that 'knowledge' does not behave like a natural kind term, and in Grundmann 2008, pp. 560–1.

[8] One might worry here that paradigm cases should be selected pre-theoretically, whereas sceptics and Mooreans rely at least on implicit epistemic theories. However, naturalists explicitly claim that intuitive classifications of things are always theory-mediated. Kornblith 1998, e.g., contends: 'Background knowledge will play a substantial role in determining a first pass categorization of samples. Judgments about which features of the rocks are even deemed relevant in classification […] are themselves theory-mediated, although the operation of theory here is unself-conscious and is better revealed by patterns of salience than it is by overt appeal to principle.' (p. 134).

contexts.[9] The process of explication follows certain rules: it should render the pre-theoretic *explicandum* more exact, theoretically more fruitful with respect to its intended use, definitionally simple, and the resulting *explicatum* should be sufficiently similar to the *explicandum*. Although Carnap originally designed the method of explication as a means to adjust folk concepts to scientific needs only, one can easily adopt it to the context of epistemology. Here we start with epistemic folk concepts and then render them sufficiently precise, coherent, definitionally simple, and adjusted to the needs of epistemology. Obviously, this enterprise of transforming a given epistemic concept is not dictated by epistemic intuitions or the competent use of folk concepts.

Explications may help us to overcome problems from cross-cultural diversity and variation. Suppose the Universality Thesis is false and hence concepts of knowledge or justified belief are not exactly co-extensional across most or all human languages. As long as there is significant cross-linguistic overlap among the epistemic terms, we may explicate these terms in the direction of their common core. In such a way, explication may serve as a means to establish a stable foundation for doing epistemology. However, the basic idea of explication is to enhance a given concept by making it more precise, coherent, and properly usable, rather than to change its content radically. It is still meant to fix, repair, and clean up a given concept. Or, as Carnap (1950, p. 7) reminds us: 'The explicatum is to be similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used…'. Consequently, the extent to which the required transformation can change the extension of the folk concept is limited. Explication is always conservative in this respect. But then it does not have the resources to overcome the challenge from insignificance in general. If the folk concepts do not refer to the most relevant and valuable things in the epistemic domain, small semantic changes that are within reach of explication need not alleviate the fundamental problem.

Conceptual ethics can be understood as a radically liberalized version of explication. Accordingly, normative conceptual choices (answers to questions like 'Shall we continue to use concept C?' or 'How shall we use C in the future?') are not motivated solely by scientific demands but rather by a plurality of values and interests that may be either domain-specific or more broadly pragmatic, moral, or even political (Burgess and Plunkett 2013a, 2013b). Moreover, the extent to which the use of a given concept can be changed is not narrowly limited. There can be pragmatic reasons to continue using numeric concepts in mathematics or physics even though it has turned out that numbers do not exist. Moreover, terms can be redefined in radical ways if this serves our political interests and goals. In this vein, Haslanger (2012) has defined 'woman' as someone who is socially subordinated in some way on the basis of presumed female gender.

Conceptual ethics fares better than orthodox explication with respect to the challenge from insignificance. Suppose that the epistemic folk terms do not capture what is epistemically most important and valuable. Then minor conceptual changes that are within reach of explications might not be radical enough to adjust the folk concepts to the relevant epistemic properties. In contrast, conceptual ethics is not restricted in the same way and may radically depart from the ordinary usage of terms, and hence permits that the epistemic folk terms are wholly transformed in light of one's epistemic priorities. The problem with

---

[9] Thanks to Steffen Koch for deepening my understanding of this point.

conceptual ethics is that it is too liberal. According to it, conceptual decisions about epistemic terms need not be governed by the goal of representing epistemic values accurately. Many non-epistemic goals, values, and interests may also rationally influence our conceptual choices in the epistemic domain.[10]

Let us finally turn towards the **epistemic desiderata approach** (Alston 2005). According to Alston, persistent and intractable controversies in epistemology indicate that the prevailing disputes are of a merely verbal rather than genuine[11] and substantial kind.[12] In his view, it is not true that at most one side has the correct understanding of justification. Rather, different epistemologists emphasize different epistemically valuable properties under the same label. In Alston's view, there are many different and irreducible properties of beliefs that possess a positive value in the cognitive domain. It is thus a virtue to cultivate 'a radically pluralistic approach' to epistemic values (Alston 2005, p. 39). Alston claims that under the overarching cognitive umbrella of achieving true beliefs, many different spheres of epistemic value reside. Among them are truth-conductiveness, discrimination of true beliefs, responsibility, or more broadly construed aims such as understanding, coherence, or systematicity.

Radical pluralism has the advantage of being integrative rather than exclusive. However, it also has a clear downside. If there are many unrelated epistemic values, how should one prioritize them? In the end, one needs a clear answer to the question of what one should ultimately believe. In this respect, an account that says what one should believe relative to each of many uncoordinated epistemic values does not help much. I do not see how Alston could respond to this coordination problem.[13]

3. Alethic Instrumentalism: The Foundations

Given the rather dim prospects of the alternatives to SAE that have been considered, let me start afresh. In order to capture what really matters epistemologically, we should set aside our epistemic folk terms and begin with an examination of the final (purely) epistemic goal of cognitive inquiry. Further epistemic values will then be merely instrumental with respect to this goal. Whether any given method or cognitive process is instrumentally valuable in this respect is a question that cannot be settled by philosophers alone.[14] However, epistemology can outline the basic requirements for instrumentally valuable methods. Moreover, epistemology can integrate these requirements into a comprehensive theory and then construct

---

[10] Admittedly, my own approach might be understood as a specific kind of conceptual ethics such that conceptual choices are only guided by purely epistemic goals and values.

[11] In a *genuine* dispute the opposing parties do not talk past each other. Nevertheless, in such a dispute both parties may be right, if one regards truth-relativism as a serious option. In contrast, *substantial* disputes are such that at most one party gets it right.

[12] Alston's argument is feeble. First, it over-generalizes. Since there is no convergence with respect to the vast majority of philosophical views (and not only in epistemology), philosophers would talk past each other all the time. That is hard to believe. Second, one need not invoke verbal disputes to explain the persistence of controversies: they may also be the result of underlying, unnoticed methodological differences. See Grundmann (forthcoming).

[13] Joachim Horvath has suggested to me that Alston might take a subjectivist stance here. He then would argue that each subject prioritizes the values that fit best with her overall preferences. However, this would relativise epistemology to individual agents.

[14] Often, the question of how reliable certain instruments and heuristics are is an empirical matter.

new technical terms that properly reflect the relevant epistemic desiderata. This approach will be called 'alethic instrumentalism'. It is based on genuine epistemic values rather than on our folk concepts. Hence, alethic instrumentalism promises to avoid the problem of insignificance. It also avoids the coordination problem which dogged radical pluralism, since there is indeed only one final value in epistemology and all the other values are, as instrumental values, derivative or subordinated. The result will be somewhat revisionary with respect to folk epistemology.

What is the final epistemic goal? We are a bit sloppy when we usually claim that *truth* is this goal. What we are really after in our inquiries is acquiring as many true beliefs and avoiding as many false beliefs as possible (see Alston 2005).[15] So, we pursue two goals at the same time. Being ignorant of truths that are out there is an epistemically bad thing; but believing something falsely is also cognitively bad. Formally, it is possible to combine these two final goals into a single one, namely believing something if and only if it is true (Alston 2005, Ch. 2). But whenever we can only approximate this goal, it turns out that there is a trade-off between its two aspects: we can achieve more true beliefs at the cost of also acquiring some false ones, or we can avoid errors at the cost of believing only very few truths (if any).

From an epistemic point of view, both sub-goals are equal. However, practical concerns can motivate us to privilege one over the other. Suppose that it is very important to select all people who suffer from a specific illness in order to put them in quarantine. In this case, it would be wise to run the risk of misidentifying some healthy people as ill just to make sure that no ill person is left out. It seems more important to detect everybody who is in fact ill than to avoid believing falsely of some people that they are ill. In contrast, suppose that a medical diagnosis says that you suffer from an illness that can only be cured by life-threatening surgery: in that case, it seems wise to give preference to avoiding errors in the diagnosis. Both cases exhibit practical considerations that are admittedly important. But these considerations should not play any significant role from a purely epistemic point of view.

It is often claimed that only *significant* or *important* truths are the proper goals of inquiry (Bishop and Trout 2005, p. 93; Alston 2005). It is certainly true that we find only those truths *interesting* that serve our preferences or that we take to be objectively relevant. But this has relevance only for the motivation of our related projects and actions. From a purely epistemic point of view all truths should count as equal. Since resources are often scarce and cognitive projects are time-consuming, we sometimes have to prioritize cognitive projects. But these choices are made on the basis of non-epistemic reasons. What about truths that, when known, have the potential to generate further true beliefs? What I have in mind here are general principles, laws, or methodological truths—do these not deserve a prominent status as epistemic goals? I do not think so: because what is at issue here are the *final* goals of inquiry. Having true beliefs about principles, laws, or methods may be instrumentally valuable with respect to promoting the acquisition of further true beliefs, but this does not license giving it the special status of a final epistemic goal.

---

[15] Admittedly, this claim is controversial. In contrast to the view defended here, some epistemologists claim that knowledge or understanding are final values. Still others claim that epistemology has to be based on fundamental epistemic norms rather than values.

One further question naturally arises: Given that acquiring as many true beliefs and avoiding as many false beliefs as possible is our final epistemic goal or twin-goal, what determines that it is? There are three different means of explaining this cognitive goal: (i) the goal may be grounded in facts about our cognitive perspective or our cognitive system, (ii) it might be constitutive of taking a certain stance ('the epistemic point of view'), or (iii) it might be grounded in objective epistemic values. Proponents of (i) typically claim that humans (and other agents) are generally interested in truth and that this interest is even a presupposition of any successful behaviour or navigation through the world (see Kornblith 2002, pp. 137–61). Such a general human interest is said to determine the goal of cognitive inquiry. Let me raise two objections to this view. First, not any human desire or interest can fix our common *epistemic* goal—for example, the desire for happiness does not seem adequate for doing this job. Rather, we must first select the *epistemically* relevant human desires. However, this can only be done by relying on our epistemic concepts. Hence, this view is a further variant of SAE. Second, even if the human interest in truth is not completely arbitrary but is deeply embedded in our cognitive lives, there are still occasions on which we explicitly do not want to know particular truths (when knowing the truth is too cruel or when it would spoil the suspense of an interesting story). Even in these cases we can assess beliefs from an epistemic point of view that would be inadmissible to us if that point of view were constituted by existing desires in truth. A variant of (i) can handle this second objection: we need only modify it to claim that our cognitive goal is determined by the evolutionary function of our cognitive system, rather than our contingent desires (Graham 2014). Accordingly, it is the biological and evolutionarily acquired function of our cognitive system to aim at truth, as it is the biological function of the heart to pump blood through the circulatory system. Yet although this evolutionary determination of the cognitive goal makes it more robust than the desires we happen to have, it is still not robust enough. Imagine that a human-like creature, 'swampman', happens to come into existence through a lightning that strikes a swamp. This creature will lack any normative functions since it has no evolutionary history. Nevertheless, it seems perfectly possible to assess its beliefs (if it has any) from an epistemic point of view.

According to the second option (Alston 2005), the cognitive goal is not grounded in any worldly facts but is simply the result of assessing beliefs from a certain perspective that commits one to nothing much. When we take up the epistemic point of view we assess the beliefs with respect to the stipulated truth-goal. We then try to figure out how well our beliefs and methods perform *given* this goal. One can assess the instrumental value of things with respect to any stipulated goal, even if no one actually pursues this goal. It is still true that certain means are good means with respect to certain goals, no matter whether anybody actually has these goals. One advantage of this view is that it is ontologically cheap, and free of any commitments to existing goals or values. But the view has also certain disadvantages. Of course, one can judge things (beliefs) against any kind of hypothetical measure. The question remains: Why do we choose this rather than any other stance? Moreover, why do we call this stance 'the epistemic point of view'? As far as I can see, the character of the *epistemic* point of view must be fixed by our epistemic concepts. Ultimately, it is again conceptual competence that determines the starting point. Hence, all the problems of SAE and from cross-linguistic diversity recur.

I therefore think that the best strategy is to assume that the epistemic goal is grounded in the intrinsic value of having as many true beliefs and avoiding as many false beliefs as possible. Accordingly, objective facts about values determine the epistemic point of view. These value-facts make certain things desirable, even if they are not actually desired. These facts are not grounded in facts about the practical utility of having true beliefs. It is simply the fact that truth matters in itself. I cannot argue for this view here, but I take it that it is the best way of explaining the epistemic goal.

If having as many true beliefs and avoiding as many false beliefs as possible is the final epistemic goal, then further instrumental values can be derived. The proper targets of assessment are *belief-producing methods*, such as our reliance on specific instruments like calculators or speedometers, our use of methodological strategies and heuristics like relying on experts or using certain algorithms for calculation, or our innate sources of belief like perception, reasoning, or memory. When we assess the instrumental value of these belief-producing mechanisms with respect to the twin-goal of truth, we must carefully pay tribute to both aspects of this goal (maximizing true *and* minimizing false beliefs). Accordingly, a method is (instrumentally) good only if it has a positive truth-ratio, that is, it produces more true than false beliefs. This can be measured by the method's degree of reliability. Methods can be positively correlated with truth, that is to say substantially reliable ($R > 0.5$), they can be indifferent ($R \approx 0.5$), or they can be anti-reliable ($R < 0.5$). For example, reliance on tossing a coin is indifferent; reliance on counter-induction is anti-reliable. Being substantially reliable is necessary but not sufficient for performing well with respect to the overall epistemic goal. Suppose that a given method is highly reliable at the cost of producing only very few beliefs. Although the ratio of true to false beliefs is excellent, this method will not produce many true beliefs—similar, for example, to Descartes' rule of believing only what is certain. However, the cognitive goal does not only require substantial reliability, but also calls for a process to be informative, namely to generate many true beliefs. Typically, there is a trade-off between being reliable and being informative. Methods that approximate perfect reliability are often not very informative; they tell us very little about reality. Very informative sources are often not particularly reliable but generate some noise. However, sometimes it is the very fact that a method is informative that increases its reliability (Grundmann 2010). Suppose that a method M is substantially reliable but is not flawless. When M has produced false beliefs, more information provided by M itself might enable us to detect and correct these false beliefs. For example, vision does not only represent objects and scenarios, it often also represents the quality of its representations in given contexts: you can visually experience that the lighting conditions are bad or that the object is too far away (Weinberg 2007). When we discover that a belief has been formed under sufficiently unfavourable conditions, we suspend judgment. In other cases, more information about a particular target object permits calibration through coherence considerations. In any case, the assessment of cognitive methods has to take into account both reliability *and* informativeness.

Before closing this section, let me address two further issues about the foundations of alethic instrumentalism. First, one might think that deductive reasoning based on the output of substantially reliable methods always constitutes a further instrumentally good epistemic method. The thought is this: deduction is truth-preserving and therefore perfectly reliable, and it also respects the desideratum of informativeness since it is a cheap way of generating

further information. However, as the conjunction fallacy shows, for example, competent deduction need not always lead from reliably based premises to reliably formed conclusions. If the premises are not based on perfectly reliable methods, deductive reasoning with multiple premises may lead to conclusions that are based on methods with a reliability < 0.5.[16] So deductive reasoning does not always extend our instrumentally valuable methods. I will come back to this in the next section.

Second, belief-forming methods have been assessed as being instrumentally valuable with respect to the epistemic goal if they are *in fact* reliable and informative. Nothing that has been said so far supports the idea that good epistemic methods need to be validated from the epistemic agent's own perspective such that she is able to make reasonable epistemic choices with respect to the adopted methods. Such choices make sense only if (i) the assessment of a given method is non-circular and (ii) the method constitutes a general policy or strategy. Only methods that satisfy both conditions can be rationally chosen. These conditions would significantly constrain the set of admissible methods.

One standard objection against this whole idea of cognitive guidance from the agent's point of view says that doxastic voluntarism is false because beliefs do not result from epistemic choices. But no matter whether we have choices with respect to our beliefs, we certainly make choices about whether we should use a specific instrument or adopt a particular cognitive strategy.

Can we assign a particular epistemic value to cognitive methods that facilitate such epistemic choices? It is tempting to believe that critical self-regulation of methods is reliability-enhancing. The thought is that a person who permanently adjusts her adopted methods under critical reflection will be generally more reliable than someone who is equipped with a fixed stock of methods. If this is true, then methods that facilitate assessment from the first-person perspective are more instrumentally valuable with respect to the epistemic goal. However, it is highly dubious whether critical self-assessment is really reliability-enhancing (see Kornblith 2012 for a negative verdict). Our self-assessment seems to be strongly influenced by biases such as overconfidence or confirmation-bias. If this negative verdict turns out right, we cannot derive, purely from considerations of the epistemic goal, the instrumental value of methods that are reflectively assessable. Strictly speaking, methods with this characteristic seem to be preferable only for creatures like us who typically make reflective choices.

To wrap up this section: having as many true beliefs and avoiding as many false beliefs as possible is our (general) epistemic goal. This twin-goal is grounded in the objective value of true beliefs and the objective value of the absence of false beliefs. Given this goal, the instrumental value of belief-forming mechanisms depends on how reliable and how informative they are. For reflective creatures like us these methods are typically only adoptable and sustainable if their value is reflectively assessable.

4. Alethic Instrumentalism: The Constructive Part

---

[16] Hereafter, whenever I talk of reliable premises, conclusions, or beliefs, I mean—strictly speaking—premises, conclusions, or beliefs that produced by a reliable process.

We are now in a position to raise two further questions:

Q1 How can the instrumental values be combined in such a way that the methods licensed by them are epistemically minimally acceptable, given the epistemic goal?

Q2 What are the epistemically best choices among the epistemically acceptable methods?

Let me first address Q1. What would a minimally acceptable belief-forming method M look like?

1. M's reliability should be clearly above 0.5, but it need not approximate 1. M's reliability should be assessed relative to the *actual* world.[17]

The motivation for (1) is straightforward: the method must be substantially reliable if it is to promote the epistemic goal of having more true than false beliefs. But the degree of reliability should not be excessive, since that would threaten the method's informativity. Moreover, reliability in the actual world is what matters. Given that truths in the actual world are typically the goal,[18] hyper-reliability across possible worlds is not needed and reliability relative to particular other (e.g. normal) worlds is irrelevant.[19]

2. If M is not innate but must be selected by the epistemic agent, M should constitute some general rule or policy that can be assessed in a non-circular way.

(2) is motivated primarily by the contingent fact that humans are creatures that need reasons to adopt a new method. Hence, they select such methods context-independently on the basis of believing that the method is instrumentally valuable in general. (2) is not a general *epistemic* desideratum since it is not epistemically motivated, but is rather motivated by special features of human cognition.

Is this feature of methods compatible with their objective reliability? One might worry that it is not, since it sounds like a typical internalist constraint. But this is misleading. Epistemic internalism with respect to a particular epistemic property E claims either (i) that E supervenes on non-factive mental states of the cognizer,[20] or (ii) that E is reflexively accessible (i.e. by introspection plus a priori reasoning alone).[21] In contrast, (2) requires only that we can decide independently of a given method M whether it is instrumentally sufficiently good. In our deliberation we may rely on any of our innate methods of belief-formation, including perception.

3. An unrestricted principle of closure under competent deduction does not hold for the extension of minimally acceptable methods.

---

[17] Strictly speaking, we need more than that: we need reliability in the actual world over time. Otherwise, we could not rule out methods that when used once lead to a true belief as unreliable. Thanks to Chris Ranalli for discussion about this point.

[18] Here I leave aside the issue of modal truths.

[19] In contrast, Goldman 1986 and Graham 2016 have both argued that it is not reliability in the actual world but reliability in normal worlds that matters for epistemic evaluation.

[20] This is what mentalism claims.

[21] This is what access internalism claims.

If all minimally acceptable methods were perfectly reliable (R = 1), deductive reasoning on the basis of reliable premises would always constitute a further reliable method. However, minimally acceptable methods need not be perfectly reliable because otherwise they would lack informativity (see (1)). But if minimally acceptable methods are sufficiently rather than perfectly reliable, the principle of closure does not hold in all cases of multiple-premise deductions. For example, even if we can reliably believe of each ticket of a fair lottery that it will lose, we cannot reliably believe that all tickets will lose. Hence the price of informative methods is the restriction of closure.

To avoid confusion, the constructive part of alethic instrumentalism requires some new terminology. Epistemically minimally acceptable methods will be called 'epistemically competent methods', and beliefs that are produced by epistemically competent methods are called 'epistemically competent beliefs'.

Let us now turn to Q2. In a particular epistemic situation, we want to find out whether a particular proposition p is true. Typically, many different competent methods that are informative with respect to p are available to us. What is the best choice for us in this situation? It seems obvious that we should choose the competent method that is most reliable. When the target proposition is fixed, no further considerations play any role for our selection among the competent methods that are informative with respect to the target proposition. Since the best choice from this stock will typically not be perfectly reliable, the output of the best choice need not be true. Methods that are epistemic best choices in a particular situation will be called 'epistemically optimal methods'; beliefs that are produced by epistemically optimal methods will be called 'epistemically optimal beliefs'.

In some respects, the resulting view seems to be rather distant from folk epistemology. For a number of reasons, what English speakers call 'knowledge' does not play any significant role within alethic instrumentalism. First, our general assessment of competent or optimal methods is not sensitive to highly situational (un-)Gettierizing factors. Second, methods that are perfectly reliable in general and hence produce knowledge in the folk sense are too rare to be a main topic in epistemology. Third, knowledge also cannot be analysed as epistemically optimal true belief, since the combination of truth and being produced by an optimal method is not interesting from a theoretical point of view. Moreover, in alethic instrumentalism there is no place for any internalist conception of justification, nor for epistemic norms. This new framework deviates radically from folk epistemology. To avoid confusion, it seems wise to use the new terminology introduced above.

5. Objections

Let me finally address two obvious objections to alethic instrumentalism. Here is the first: Suppose that an epistemic agent A produces the belief that p on the basis of a competent (i.e. sufficiently reliable and informative) method M. Further suppose that there is a body of (confirming or disconfirming) evidence with respect to p that is not incorporated in M. From the perspective of alethic instrumentalism this fact does not seem relevant to the epistemic assessment of A's belief that p. However, disputing the epistemological relevance of further evidence seems deeply mistaken.

Reply: Having used a competent method in forming the belief that *p* does not exhaust the data that are epistemically relevant to alethic instrumentalism. If the actually operative method ignores some available evidence, there is probably another method available that makes use of this further evidence and is therefore more reliable. Hence, A's belief that *p* is not epistemically optimal unless it is based on a method that incorporates all available evidence.

Second objection: In justifying the new framework of epistemic evaluation I rely on the old conceptual framework that uses notions such as 'reason', 'evidence', 'justification', or 'knowledge'. However, by replacing the old framework I denigrate the very reasons I had for adopting the new framework. Hence, when I claim that the old framework of justification should be given up I undermine my reasons for doing that very thing. The revisionary approach to epistemology seems to be self-undermining.

Reply: The fact that the new framework can be defended within the old framework does not imply that the new framework can *only* be defended in the old one. Since the new framework is not inconsistent with the old one, it is possible (and in fact probable) that it is also optimal by its own standards.

6. Conclusion

There is a general worry that *Standard Analytic Epistemology* rests on a fundamental mistake, massively overemphasizing the role that epistemic folk concepts such as 'justification' or 'knowledge' play within epistemology. Epistemic folk concepts may not be a proper basis for stable epistemic intuitions, they may be cross-culturally diverse, and they may missrepresent what is epistemically important. Although it is not yet fully clear how serious these concerns are, it is high time to try out an approach to epistemology that is not based on our folk concepts. One promising approach in this direction is what I call 'alethic instrumentalism'. We can understand this as a kind of truth-technology that assesses the instrumental value of belief-forming mechanisms with respect to the properly qualified truth-goal. This approach has much in common with what Bishop and Trout (2005) call 'strategic reliabilism'. However, it is stricter in developing a purely epistemic assessment that does not involve any pragmatic elements. Although this new approach reflects some important features of our old epistemological framework, it wipes out others altogether. From the perspective of alethic instrumentalism, internalist aspects of justification, non-actual world reliability, unrestricted closure, or knowledge do not play any significant role. To avoid confusion, this radical change in perspective should be highlighted by a change in epistemic terminology. The next step would be an assessment of real-life methods, heuristics, and processes that rely on this new framework.

References

Alexander, Joshua 2012: *Experimental Philosophy: An Introduction*. Cambridge, UK/Malden, MA: Polity Press.

Alston, William 2005: *Beyond 'Justification'. Dimensions of Epistemic Evaluation*. Cornell.

Beckermann, Ansgar 2001: 'Zur Inkohärenz und Irrelevanz des Wissensbegriffs. Plädoyer für eine neue Agenda in der Erkenntnistheorie', in: *Zeitschrift für philosophische Forschung* 55: 571–93.

Bishop, Michael and J. D. Trout 2005: *Epistemology and the Psychology of Human Judgment*. Oxford: Oxford University Press.

Burgess, Alexis and David Plunkett 2013a: 'Conceptual Ethics I', in: *Philosophy Compass* 8: 1091–1101.

Burgess, Alexis and David Plunkett 2013b: 'Conceptual Ethics II', in: *Philosophy Compass* 8: 1102–1110.

Carnap, Rudolf 1950: *Logical Foundations of Probability*. Routledge.

Craig, Edward 1990: *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Oxford University Press.

Fantl, Jeremy and Matthew McGrath 2009: *Knowledge in an Uncertain World*. Oxford: Oxford University Press.

Fassio, Davide and Robert McKenna 2015: 'Revisionary Epistemology', in: *Inquiry* 56: 755–79.

Goldman, Alvin 1986: *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.

Goldman, Alvin 1988: 'Strong and Weak Justification', in: *Philosophical Perspectives* 2: 51–69.

Goldman, Alvin 1999: *Knowledge in a Social World*. Oxford: Oxford University Press.

Goldman, Alvin 2007: 'Philosophical Intuitions: Their Target, Their Source, and Their Epistemic Status', in: *Grazer Philosophische Studien* 74: 1-26.

Graham, Peter 2014: 'Functions, Warrant, History', in: Abrol Fairweather and Owen Flanagan (eds.), *Naturalizing Epistemic Virtue*, Cambridge University Press, pp. 15–35.

Graham, Peter 2016: 'Against Actual-World Reliabilism: Epistemically Correct Procedures and Reliably True Outcomes', in: Miguel Angel Fernandez (ed.), *Performance Epistemology: Foundations and Applications*, Oxford: Oxford University Press.

Grundmann, Thomas and Joachim Horvath 2014: 'Thought Experiments and the Problem of Deviant Realizations', in: *Philosophical Studies* 170: 525–33.

Grundmann, Thomas 2008: *Analytische Einführung in die Erkenntnistheorie*, Berlin/New York: DeGruyter.

Grundmann, Thomas 2010: 'Some Hope For Intuitions. A Reply to Weinberg', in: *Philosophical Psychology* 23: 481–509.

Grundmann, Thomas 2017: 'Platonism and the Apriori in Thought Experiments', in: James Robert Brown, Yiftach Fehige and Michael T. Stuart (eds.), *The Routledge Companion of Thought Experiments*, London/New York: Routledge.

Grundmann, Thomas forthcoming: 'Progress and Historical Reflection in Philosophy', in: Marcel van Ackeren (ed.), *Philosophy and the Historical Perspective*, Proceedings of the British Academy.

Haslanger, Sally 2012: 'Gender and Race: (What) Are They? (What) Do We Want Them to Be?', in: Sally Haslanger, *Resisting Reality*, Oxford: Oxford University Press, pp. 221–47.

Hazlett, Allen 2010: 'The Myth of Factive Verbs', in: *Philosophy and Phenomenological Research* 80: 497–522.

Horvath, Joachim 2016: 'Conceptual Analysis and Natural Kinds: The Case of Knowledge', in: *Synthese* 193: 167–84.

Kim, Minsun and Yuan Yuan 2015: 'No Cross-Cultural Differences in Gettier Car Case Intuitions', in: *Episteme* 12: 355–61.

Kornblith, Hilary 1998: 'The Role of Intuition in Philosophical Inquiry: An Account with No Unnatural Ingredients', in: Michael DePaul and William Ramsey (eds.), *Retinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, Lanham: Rowman and Littlefield, pp. 129-41.

Kornblith, Hilary 2002: *Knowledge and Its Place in Nature*, Oxford: Oxford University Press.

Kornblith, Hilary 2012: *On Reflection*, Oxford: Oxford University Press.

Machery, Edouard et al. 2015: 'Gettier Across Cultures', in: *Nous* Early View, DOI: 10.1111/nous.12110.

Matthewson, Lisa and Jennifer Glougie forthcoming: 'Justification and Truth: Evidence From Languages of the World', in: Stephen Stich and Masaharu Mizumoto (eds.), *Epistemology For the Rest of the World*, Oxford: Oxford University Press.

Mizumoto, Masaharu forthcoming: '"Know" and Japanese Counterparts "shitte-iru" and "wakatte-iru"', in: Stephen Stich and Masaharu Mizumoto (eds.), *Epistemology For the Rest of the World*, Oxford: Oxford University Press.

Nagel, Jennifer, Raymond Mar, and Valerie San Juan. 2013: 'Authentic Gettier Cases: A Reply to Starmans and Friedman', in: *Cognition* 129: 666–69.

Pasnau, Robert 2013: 'Epistemology Idealized', in: *Mind* 122: 987–1021.

Putnam, Hilary 1975: 'The Meaning of "Meaning"', in: Hilary Putnam, *Mind, Language and Reality*, pp. 215–71.

Sartwell, Crispin 1992: 'Why Knowledge is Merely True Belief', in: *The Journal of Philosophy* 89: 167–180.

Seydsayamdost, Hamid 2015: 'On Normativity and Epistemic Intuitions: Failure of Replication', in: *Episteme* 12: 95–116.

Stanley, Jason 2005: *Knowledge and Practical Interests*. Oxford: Oxford University Press.

Stich, Stephen and Mizumoto, Masaharu forthcoming: 'Manifesto', in: Stephen Stich and Masaharu Mizumoto (eds.), *Epistemology For the Rest of the World*, Oxford: Oxford University Press.

Stich, Stephen 1998: 'Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity', in: Michael DePaul and William Ramsey (eds.), *Rethinking*

Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry, Lanham, pp. 95–127.

Weinberg, Jonathan 2007: 'How to Challenge Intuitions Empirically Without Risking Skepticism', in: *Midwest Studies in Philosophy* 31: 318–43.

Weinberg, Jonathan; Nichols, Shaun and Stich, Stephen 2001: 'Normativity and Epistemic Intuitions', in: *Philosophical Topics* 29: 429–60.