The epistemology of Ernest Sosa: an introduction

Christoph Kelp¹

Published online: 13 October 2020
© Springer Nature B.V. 2020

There can be no question that Ernest Sosa is one of the most influential voices in contemporary epistemology. He has made pathbreaking contributions to a wide range of topics in the field and beyond, including on its most central issues such as the nature of knowledge, the structure of knowledge, the value of knowledge and the extent of knowledge.

It is fair to say that his most widely discussed contributions, at least in recent times, are on virtue epistemology and safety conditions on knowledge. Whilst both topics are intimately related in Sosa's own work, they have generated discussions that have lives of their own. Since the bulk of the contributions to this special issue also focus on these two topics, I will take a few moments to sketch a few key ideas in what follows.

1 Safety

The core claim of a safety condition on knowledge is that knowledge requires belief that is safe from error. Very roughly, what this means is that one's belief must not only be true, but it must also be the case that one might not easily have believed falsely. Or, given a standard possible worlds semantics for 'might', to say that one believes safely is to say that one not only believes the truth at the actual world, but one also avoids error at nearby possible worlds. In this way, safety is a modal condition on knowledge.

Accounts of knowledge that countenance a safety condition on knowledge are among the most popular in contemporary epistemology. This is no accident. After all, safety-based accounts of knowledge promise to offer solutions to some of epistemology's most difficult problems, including the Gettier problem and the problem of scepticism. Very roughly, the Gettier problem is the problem of identifying the missing link between justified true belief and knowledge. That there is a missing link was first argued by Edmund Gettier in his influential 1963 paper. Gettier offers a couple

✉ Christoph Kelp
christoph.kelp@glasgow.ac.uk

¹ COGITO Epistemology Research Centre, University of Glasgow, 67-69 Oakfield Ave, Glasgow G12 8LP, UK

of cases that showed that one can have a justified true belief that isn't knowledge. To get a better grip on this, let's take a look at two famous examples:

Stopped clock You take a competent reading from a clock that you know to be reliable and have no reason to think is currently not working. Based on this reading you form a belief that it is noon. What's more, your belief is true: it is indeed noon. Crucially, however, the clock is broken and the reason your belief is true is that it happened to stop working exactly twelve hours ago.

Fake barns You are driving through the countryside and see what appears to be a barn in the field. Based on this, you form a belief that there is a barn before you. What's more, your belief is true: there is a barn before you. Crucially, however, the barn is the only real barn in an area otherwise peppered with fakes that are so cleverly constructed as to be indistinguishable from real barns from your position on the road.

In both of these cases, your beliefs fall short of knowledge. At the same time, they are both justified and true. One attractive feature of safety-based accounts of knowledge is that they can explain the absence of knowledge here. To see how, recall that safety requires avoidance of error not only at the actual world but also at nearby possible worlds. In the above cases, you do get it right at the actual world. Crucially, however, you fall into error at the vast majority of nearby possible worlds. In *Stopped Clock*, there are many nearby possible worlds at which you look at the clock at least a minute earlier or later. At those worlds, the clock still displays noon and you end up with a false belief. In *Fake Barns*, there are many nearby possible worlds at which you look out of the window of your car a little earlier or later. At many of those worlds, you will look at a fake barn and end up with a false belief. This means that your beliefs aren't safe in either of those cases. In this way, safety-based accounts of knowledge promise to explain the absence of knowledge in Gettier cases.

The perhaps most influential sceptical argument is the argument from ignorance. Here is a version of it:

AI1. You don't know that you are not a handless brain in a vat (= BIV).

AI2. If you don't know that you are not a handless BIV, then you don't know that you have hands.

AI3. Hence, you don't know that you have hands.

AI1 is highly intuitive and AI2 is motivated by the highly plausible closure principle for knowledge, according to which knowledge transmits across competent deduction. But, of course, if both AI1 and AI2 hold, the unpalatable sceptical conclusion, AI3, follows straightforwardly by modus ponens.

The perhaps most prominent alternative modal condition on knowledge, sensitivity, famously offers a response to the argument from ignorance that denies the closure principle and hence AI2. According to sensitivity, very roughly, one knows that *p* only if were *p* false, one wouldn't believe that *p*. It is easy to see that your belief that you have hands is sensitive: if you didn't have hands, say because you lost them in an accident, you wouldn't believe that you did. At the same time, your belief that you are not a handless BIV is not sensitive: if you were a handless BIV, everything would still look the same, and so you would still believe that you aren't. On sensitivity accounts

of knowledge, AI1 is true and AI3 is false. Hence, AI2 must be false as well as the closure principle that underlies it.

While its champions have touted this response as one of the attractions of sensitivity-based accounts of knowledge, there has been a growing consensus that the costs associated especially with denying closure are too high. Another important motivation for safety-based accounts is that they can offer an attractive alternative which preserves closure. In particular, they can offer a Moorean response to problem, according to which AI3 is false and AI2 is true with the result that AI1 is false. After all, your belief that you have hands is safe: there are no nearby worlds at which you falsely believe that you have hands: AI3 is false. Similarly, there are no nearby worlds at which you are a handless BIV and hence no nearby worlds at which you falsely believe that you are not a handless BIV: AI1 is false. In this way, safety-based accounts of knowledge promise not only to allow us to hold on to closure but also to secure a robust anti-sceptical solution to the problem of scepticism.

2 Virtue epistemology

Central to Sosa's virtue epistemology – and, as a result, to much of contemporary virtue epistemology in the reliabilist camp – is a normative framework for the assessment of performances with an aim. In the first instance, performances with an aim can be evaluated along the following three dimensions: accuracy, adroitness and aptness. Roughly, a performance is accurate if and only if it attains its aim, it is adroit if and only if it is produced by the exercise of a competence to attain its aim, and it is apt if and only if it is accurate because adroit. For instance, your shot in basketball is accurate if and only if it goes in, it is adroit if and only if it is produced by an exercise of a competence to make shots, and it is apt if and only if it goes in because of the exercise of the competence to make shots.

While this gives us the basic account for first order evaluations of performances, Sosa does not take this to be the whole story about the normativity of performances. Rather, he countenances two further types of aptness, alongside first order aptness, or 'animal' aptness as Sosa calls it. These additional types of aptness are 'reflective' and 'full' aptness. Attaining these further types of aptness requires accurate and indeed apt performance at a higher order, in addition to animal aptness. In a nutshell, the thought is that performances will rise to these higher levels of aptness only if, alongside animal aptness, one has aptly ascertained that one's performance is free from any relevant risk one may be running: one must have arrived at an apt awareness that one's performance would be apt. While animal aptness in conjunction with apt risk assessment will be enough for reflective aptness, full aptness additionally requires that first and second order aptness are connected in the right way: one must be guided to animal aptness by one's reflectively apt risk assessment.

There are a number of normative properties that performances can enjoy. Crucially, according to Sosa, full aptness enjoys special status among these properties. More specifically, according to Sosa, full aptness is the fully desirable status for performances in a given domain; performances fall short unless they attain full aptness.

Moreover, he is also clear that this claim holds with full generality. In view of this, it seems fair to attribute the following thesis to Sosa:

The Normative Thesis (NT) Any performance in any domain attains fully desirable status qua performance in that domain if and only if it is fully apt; and it falls short qua performance in that domain if and only if it isn't.

Why think NT is true? To see how Sosa ventures to answer this question, let's return to the basketball case. In particular, let's look at two versions of it. In both versions (i) your shot is animal apt and (ii) you are very close to the end of the range of your competence to make shots. Now, in the first version of the case, you are blissfully ignorant of just how far your range extends. As a result, had you been a step further away, you would have taken the shot just the same. In contrast, in the second version of the case, you are well aware of how far your range extends. In particular, you are aware that while you are just within your range, any further away from the basket would already be too far for you. In view of this awareness of your range, you take the shot. In this case, had you been a step further away, you wouldn't have taken the shot.

Sosa's observation here is that, in the first version of the case, your performance falls short qua basketball performance. You may rightfully be scolded by your coach for ignoring the importance of shot selection. Only in the second version will you remain clear of such criticism. Since your performance is animal apt in both versions of the case, fully satisfactory performance must require more than animal aptness. Sosa's proposal is that what is required in addition is apt performance at the second order and, in particular, full aptness.

With these points about Sosa's general account of the normativity of performances in play, let's move on to epistemology. In order to connect these two issues, Sosa embraces the following thesis:

The Psychological Thesis (PT) Various psychological categories – most importantly, guessing, belief and judgment – are species of affirmation and, as a result, performances with an aim. Sosa's main interest is with affirmations with a specifically epistemic aim, which, at a minimum, involves truth.

While Sosa countenances a variety of psychological categories with epistemic aims, his main focus is on judgment (and judgmental belief). Judgment differs from other psychological categories in that it has a particularly robust epistemic aim: judgment aims not only at truth but at aptness. For a judgment to be apt, then, more is required than merely apt affirmation. What is needed for apt judgement is that one is guided to aptness by apt risk assessment. An apt judgement is a fully apt affirmation.

Finally, the last key thesis of Sosa's view is epistemological in nature. Here goes:

The Epistemological Thesis (ET) Human knowledge is apt judgment.

The attractions of Sosa's framework are not hard to discover. As a first observation, note it can offer a highly attractive account of the value of human knowledge. After all, by PT, apt judgment is fully apt affirmation. By NT, fully apt performances have fully desirable status, and so apt judgment enjoys special normative status. Finally, by ET, apt judgment is knowledge. As a result, knowledge enjoys special normative status, which, in turn, goes a long way toward explaining the special value of knowledge.

Another important motivation for Sosa's framework is that it also offers a promising solution to the Gettier problem. It may be worth noting that the solution differs from the solution offered by safety-based accounts of knowledge. No surprise, then, that the two ideas have generated their own discussions. In a nutshell, agents in standard Gettier cases—e.g. *Stopped Clock*—don't even rise to the level of animal aptness. Very roughly, here is why. We find cases with the structure of Gettier cases outside of the epistemic domain. For instance, suppose that you take a shot in basketball. While your shot is both accurate and adroit, on its way to the target, it is brought off its trajectory by a gust of wind. Crucially, however, you have a helper with a wind machine in the wings who sees to it that your shot gets back on its original trajectory and finds the target. Here your shot is accurate and adroit but not accurate because adroit and hence not apt. Standard Gettier cases share the same structure with this kind of case. For instance, in *Stopped Clock*, the fact that the clock you are taking a reading from is stopped brings your affirmation off its path to the truth, as it were. The fact that you look at the clock exactly twelve hours after it broke brings it back on target. While your affirmation is accurate and adroit, it is not accurate because adroit. Hence, it is not apt and so is bound to fall short of knowledge.

Fake barn cases are a little more difficult to handle for Sosa's framework. The reason for this is that non-epistemic analogues of fake barn cases are apt performances. For instance, in a case in which you are shooting at the only basket in the area that isn't sabotaged to prevent any shot from going in, your shots may still be apt. By the same token, there is pressure on Sosa to acknowledge that your affirmations in Fake Barns and similar cases are apt also. And, indeed, Sosa does grant this much. He explains the absence of knowledge in these cases in terms of a lack of apt judgment, where apt judgment requires fully apt affirmation. Crucially, while your affirmation in Fake Barns may be apt, it is not fully apt. After all, you get the risk of error wrong in this case: you take it to be low, while in fact it really is high. Since fully apt affirmation requires an apt risk assessment, it follows that your affirmation is not fully apt. As a result, your judgment in this case isn't apt and so falls short of knowledge after all.

3 The contributions

This special issue features sixteen novel contributions that engage critically or constructively with Sosa's epistemology. The majority of them focus either on issues relating to safety-based accounts of knowledge or his virtue epistemological framework.

Among the contributions on issues relating to safety is Sven Bernecker's 'Against Global Method Safety' which critically engages with the global method safety account of knowledge. Bernecker argues that global method safety is too restrictive because it cannot account for cases where there is a fine-grained belief that is unsafe and a relevantly similar coarse-grained belief that is safe and where both beliefs are based on the same method.

John Greco's 'Safety in Sosa' focuses on the relationship between virtue and safety. Greco argues that Sosa's positions in *A Virtue Epistemology* and in *Judgment and Agency* regarding this question are, despite appearances to the contrary, in fact not

only consistent but also well motivated in that his virtue epistemology explains why knowledge should require apt belief, and why aptness should imply safety.

Thomas Grundmann's 'Saving Safety from Counterexamples' develops a specific variant of method-relativized safety and uses it to offer a comprehensive defense of the safety account of knowledge. More specifically, Grundmann looks at a variety of putative counterexamples to safety and argues that none of them is ultimately successful.

José Zalabardo's 'Safety, Sensitivity and Differential Support' argues against Sosa's claim that sensitivity cannot be differentially supported over safety as the right requirement for knowledge. His main contention is that, although all sensitive beliefs that should be counted as knowledge are also safe, some insensitive true beliefs that shouldn't be counted as knowledge are nevertheless safe.

The majority of the contributions to this special issue focus on Sosa's virtue epistemology. One noteworthy subset addresses his claim that a range of psychological categories can be identified with species of affirmation, including guessing, belief and judgment, and its relevance to epistemology. Adam Carter's 'Sosa on Knowledge, Judgment, and Guessing' is a case in point. Carter argues that there is an interesting disanalogy between the Sosa's examples of performances in the epistemic and the practical domains, which is sourced in a problematic picture of the cognitive performance of guessing, and he develops a more promising alternative.

According to one important objection to Sosa's framework, beliefs are states, which means that they are not dynamic or telic like performances. Matthew Chrisman's 'Performance Normativity and Here-and-Now Doxastic Agency', engages with a natural response which argues that belief *formation* rather than belief itself is the proper target of epistemic normativity. Chrisman rejects this response here on grounds that obscures the "here and now" exercise of cognitive agency that he argues is central to any account of epistemic normativity and doxastic agency.

Kate Nolfi's 'Functional Belief and Judgmental Belief' explores whether we should accept Sosa's distinction between mere functional belief and judgmental belief, and, if we should, how recognizing this distinction ought to shape our epistemological theorizing. Nolfi develops series of concerns regarding whether Sosa's division between functional belief and judgmental belief is well-founded, and explores whether a virtue-theoretic performance epistemology ought to embrace the sort of two-tiered account of cognitive performance that Sosa favours.

Another noteworthy subset of contributions to Sosa's virtue epistemology concerns his account of the normativity of performances, and its relation to epistemology. Paul Horwich's 'Sosa's Theory of Knowledge' runs through the individually necessary and jointly sufficient conditions for knowledge that Sosa proposes. Horwich elaborates these conditions a little, exploring in a sympathetic spirit his reasons for including them and finally proceeds to some questions and reservations about general matters of methodology, motivation, and philosophical import.

Christoph Kelp, Cameron Boulton, Fernando Broncano-Berrocal, Paul Dimmock, Harmen Ghijzen and Mona Simion's 'Hoops and Barns: A New Dilemma for Sosa' critically assesses Sosa's normative framework for performances as well as its application to epistemology. They argue Sosa must either construe full aptness as strong enough to get barn façade cases right, but then full aptness isn't required for perfor-

mances to be fully satisfactory. Or else he construes full aptness as weak enough to avoid this problem but then he will not be able to deal with barn façade cases in the way envisaged.

Matt McGrath's 'Sosa on Epistemic Value: A Kantian Obstacle' argues, *pace* Sosa, that the quality or worth of an attempt is independent of whether it succeeds. In particular, an attempt can be fully worthy despite being a failure. He considers an alternative that places the notion of adroitness (rather than accuracy) centre stage and argues that it remains unsatisfactory. Finally, McGrath suggests the notion of a belief's fit with what the subject has to go on, a notion missing from Sosa's competence-theoretic framework, is crucial to explaining epistemic worth.

Kurt Sylvan's 'Can Performance Epistemology Explain Higher Epistemic Value' focuses on Sosa's latest effort to explain how higher epistemic value of the sort missing from an unwitting clairvoyant's beliefs might be a special case of performance normativity, with its superior value following from truisms about performance value. He argues that the new effort rests on mistaken assumptions about performance normativity. Once these mistaken assumptions are exposed, it becomes clear that higher epistemic value cannot be a mere special case of performance normativity, and its superiority cannot be guaranteed just by truisms about performance value.

Ralph Wedgwood's 'The Internalist Virtue Theory of Knowledge' defends an account of knowledge according to which for you to know a proposition *p* is for you to have an outright belief in *p* that is correct precisely because it manifests the virtue of rationality. This definition resembles Sosa's view, except that (a) on this definition, the only virtue that must be manifested (at least to some degree) in all instances of knowledge is rationality, and (b) no reductive account of rationality is attempted—rationality is assumed to be an irreducibly normative notion.

Finally, there are a number of papers that address broader issues relating to Sosa's work. Miguel Ángel Fernández Vargas's 'The Conditional Position Problem for Epistemic Externalism' develops a problem that arises for Sosa's view as a distinctively externalist epistemology. The problem is that, due to a phenomenon of epistemic circularity, one is unable to attain the reflective knowledge that one is justified in believing that perception is reliable and is confined to the merely conditional position that one is so justified if perception is reliable.

Jesper Kallestrup's 'Group Virtue Epistemology' explores an extension of Sosa's framework to a social setting in which groups constitute epistemic agents over and above their individual members. The claim is that groups can be ascribed knowledge in virtue of hitting the truth through exercising their competences in appropriate shapes and situations. Kallestrup argues that while knowledge at the collective level may diverge from knowledge at the individual level, the competences of groups are nothing over and above the combined competences of their members. In this way, the ensuing view thus has implications for the debate over reduction and supervenience in collective epistemology.

Could it be right to convict and punish defendants using only statistical evidence? Clayton Littlejohn's 'Truth, Knowledge, and the Standard of Proof in Criminal Law' takes up this question. Littlejohn argues that it is not and explains why it would be wrong. He provides reason to believe that to make proper sense of statistical evidence

in the law we need to revise some commonly held assumptions about epistemic value and defends the relevance of epistemology to this practical question.

Matthew Soteriou's 'Dreams, Agency, and Judgement' engages with Sosa's imagination model of dreaming. He develops a challenge for Sosa and offers a version of the imagination model of dreaming that can overcome that challenge. Soteriou's version of the imagination model of dreaming goes beyond what Sosa explicitly commits to when he outlines his view of dreams, however, it exploits ideas that are integral to a key theme in Sosa's recent writings on virtue reliabilism—namely his proposal that epistemic agency should be accorded a central place in that approach to knowledge, and his related proposal that agency is exercised in conscious judgement.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.