

ARTICLE

Reflection and Security

Ernest Sosa*

Rutgers University

*Email: sosaern@gmail.com

Abstract

“Reflection and Security” introduces a distinctive notion of default assumptions, one that applies in human performance domains generally, and uses that notion to extend virtue epistemology in new directions.

Keywords: Reflective knowledge; animal knowledge; skepticism; default assumptions; hinge propositions

1. Default assumptions

1.1. Human beings act in performance domains. It might be a sport, for example, such as soccer. It might be a profession, such as medicine or the law. While domains normally carry risk of failure, they can vary radically in how risk affects assessment. A pure guess in a quiz show might quite properly win the prize. More often, though, we aim at *apt* success, through corresponding competence, not just by luck. An oncologist would not make a diagnosis just by guessing.

Even when we aim at aptness rather than luck, however, that may still allow or even require proper default assumptions that are unsafe and not known to be true. Thus, it is part of the competence of night-time athletes to focus on their athletic performance, while ignoring whether the lights might fail. It is part of athletic competence to assume by default the satisfaction of background conditions like the stability of the lights.

An athlete can properly assume by default that the lights will stay on, as he considers whether to run for a fly ball, or to hit a crosscourt forehand, and he will attain *full credit* even if success is attained essentially by luck, since only by luck do the lights stay on. What is more, even while *knowing* the lights to be fragile, the athlete can still rely on the assumption that they will not fail. No matter how unsafe his performance may be, no matter how likely it may be that the lights will fail, the performance can be fully apt, even when the risk is *known* to the agent.

Compare what pilots can take for granted with no negligence. Pilots must run through their check list prior to takeoff, but they need not kick the tires. They can trust others on the airline staff to have checked the tires. Still, it is not only what pilots know with safety that they can trust without negligence to be true. For example, they might trust without negligence that no terrorism is in the offing even when it is a near certainty. Even if the terrorist will miss the flight, so that disaster will be avoided, consider the high probability of a terrorist on that flight. That is still a dire danger faced by the pilot as he prepares to take off. But no such mere danger can lower the quality of that pilot's performance, nor reduce in the slightest the credit earned for piloting the plane to its destination. Not all risks of failure bear on quality of performance.

However, the pilot's performance would *not* have retained that high quality if the pilot had *known* of the highly credible threat of imminent terrorism, so that he should have aborted the flight.

An athlete can properly assume by default that the lights will stay on, as he considers whether to run for a fly ball, or to hit a crosscourt forehand, and he will attain *full credit* even if success is attained essentially by luck, since only by luck do the lights stay on. What is more, even while *knowing* the lights to be fragile, the athlete can still rely on the assumption that they will not fail. No matter how unsafe his performance may be, no matter how likely it may be that the lights will fail, the performance can be fully apt, even when the risk is known to the agent.

1.2. By contrast, walkers outside the stadium in a dangerous neighborhood cannot assume by default the truth of that same proposition, that the lights will not fail.

Yes, if they attempt to walk harmlessly, while assuming that the lights will stay on, and if this assumption turns out by luck to be true, walkers do attain the desired success of their walk. Unlike the fielder's catch, however, unlike the winning shot in night-time tennis, such walker success is owed excessively to the luck that the lights stay on. When we walk at night in a neighborhood that turns unsafe in the dark, we cannot just assume by default that the lights will not fail. Here we do seem guilty of negligence or recklessness if we ignore that possibility, so that not even implicitly do we give it any weight. If we are credibly warned, we must heed the warning. Ignoring the warning will forfeit full credit for the walk's success.

In sharp contrast, the credit for a great catch is undiminished when the fielder ignores a warning about the fragile lights.

If the warning is ignored, the walker manifests *insufficient* competence of the sort required in order to own creditably the success of their walk. And once the safety of the walk manifests *insufficient* competence, its attainment cannot be apt (cannot be one that manifests *sufficient* competence), and is hence not properly creditable to the agent.

1.3. A default assumption is an assumption that a certain background condition is satisfied. Even when its satisfaction is required for the presence of a corresponding competence, a *background* condition can obtain *unsafely*, with no detriment to the domain-internal performance of an agent who assumes by default that it obtains, without knowing that it does. Agents sometimes risk the failure of a background condition, but running *that* sort of risk affects neither the quality of the agent's domain-internal performance, nor the credit earned for its success.

1.4. Epistemology's main domain is that of attempts to get it right aptly, along with proper storage of such beliefs. Thus, consider risk that one's epistemic competence may be diminished or lost *through failure of background conditions*. Such risk need not affect the epistemic quality of our attempts to get it right, nor the credit we earn for their success, and in particular need not affect whether we know. The risk that does affect quality and creditability of success is risk inversely proportional to degree of competence. *Such* high risk is tantamount to low competence, which surely does affect quality of performance. Creditable achievement requires success by competence, not by a fluke.

A similar conclusion may be drawn about radical skeptical scenarios. Conscientiously enough, without negligence or recklessness, we normally assume ourselves free of skeptical scenarios. And this assumption is proper even on the rare occasions when it is true but not known to be true and even quite unsafe.

1.5. Urgent, credible threats can still be harmless to quality of performance in domains *properly sealed* against external values.

What does such “proper sealing” amount to? Not that the domain just lays it down that one can properly disregard external threats *all things considered*. The point is *rather* the following.

Suppose one continues to make domain-internal attempts while dismissing all such threats and dangers, dismissing them by default assumption. One’s continuing attempts may then be *unwise* in the extreme, while still attaining the heights of domain-internal quality, unaffected by the risks incurred.

2. Two sorts of performance domains: the ludic versus the practical and non-ludic¹

2.1. Physical safety and well-being are basic, generally valued goods beyond any game or sport. Ordinary domains of praxis within which we pursue such generally valued goods should be distinguished from domains with internal constitutive goods, many of them ludic, such as the many sports that obsess and delight contemporary culture.

However, a non-ludic domain ordered through pursuit of some objective good can still be sealed off in the way of ludic domains such as baseball. Successful piloting of a plane to its destination is fully creditable even if based on a mere default assumption that the plane will not be blown up by a terrorist, an assumption that is both unsafe and not known to be true.

Another example: the domains of farming, where people aim for good crops. Actions in these domains can be sealed off, so it is not *only* ludic domains that are properly sealed off for telic assessment. The performance of a farmer can also be assessed telically in abstraction from any pragmatic factors that may also bear on its *overall* assessment.

Knowledge is a valuable human commodity, as is food. Its production is of great interest to human communities, as is the production of food. And such pursuit is properly governed by practices relevant to corresponding telic assessment, whether what we harvest is food, or, alternatively, knowledge.

2.2. Consider a fielder’s competence on a given occasion to perform well in the outfield by making catches. Such competence is determined by the athlete’s pertinent Situation, Shape, and Skill. Each of these involves the satisfaction of certain conditions. One situational requirement in baseball, for example, is that the field remain sufficiently well lit.

If the lights are nearly certain to fail imminently, a fielder’s all-out running attempt is almost sure to fail. Yet the fielder’s knowing this to be so has zero effect on the aptness and the full aptness of his great catch, provided he still acts with proper sensitivity to the factors that do matter, such as the trajectory of the ball above him, along with his tracking ability and foot speed. The fielder’s taking such factors into account is crucial to the full aptness of his attempt. Factors like these he cannot neglect to consider at least quickly and implicitly. He would be negligent if he ignored them, and might be reckless if he assigned them improper weight.

By contrast, the fragility of the lights he can ignore or dismiss with no domain-internal negligence or recklessness. His great catch is equally admired and credited, as a baseball catch, despite the lights’ being fragile, and even despite his knowing them to be fragile.

2.3. Take the risk recklessly dismissed by a basketball player who sinks a basket from the opposite end of the court when there was plenty of time to dribble or pass instead. That success (so bittersweet, certainly to the coach) will hardly earn full, unalloyed credit.

The coach will not be pleased with the play selection. But does that affect the quality of the *shot*?

¹Ludic comes from the Latin; it and its cognates apply to play, as in games or sports.

Well, the shot was poorly selected, too risky, dumb, *outrageous* really. The coach will be unhappy with the player, and specifically with those qualities of the shot itself. The shot is deplorable because poorly chosen, much worse than many easily available alternatives. *Its guiding intention is constitutive of the intentional action which is that shot itself.* And the quality of that intention is surely dependent on the agent's operative rationale. If the rationale is bad, that rubs off on the quality of that intention, and it rubs off in turn on the quality of the intentional action that is the shot itself, so that even if the shot succeeds, it will not be sufficiently through the quality of performance that guides the attempt, and it will have limited athletic worth creditable to the agent.

That player's rationale is *not* bad by being prudentially or morally or esthetically bad. Nor is it even necessarily bad *all things considered*. All these respects of badness would be irrelevant to the *athletic* quality of that shot, and more specifically to its *basketball* athletic quality. As a basketball performance, the shot is poorly selected, too risky, dumb, even outrageous. The *basketball* rationale that motivates the choice and intention to so shoot is an awful rationale, and it is the quality of that basketball rationale that rubs off on the basketball choice and intention, and rubs off in turn on the basketball action essentially constituted by that basketball choice and intention.

2.4. A distinction should be drawn here, one that will prove helpful in epistemology. We should distinguish the *selection* quality of an attempt from its *execution* quality. Take that long shot so poorly selected given the far better available basketball alternatives. That shot itself has low quality in respect of its selection. But its execution may have been superb nonetheless, in that the degree of shooting competence from that distance may have topped the scale, so that the shot's success manifested outstanding competence.

Compare a tiebreaker second serve with ad out in a tennis match tied at two sets each. Losing that service point will lose the match. Yet the server serves a flat serve at top speed, foolishly taking enormous risk. That serve is a superb ace, so that its execution quality is extremely high while its selection quality is extremely low.

Turning to epistemology, consider an inquiry into the causes of prostate cancer that, as a partial means to that objective, opts to determine the frequency of sneezing in girls who enter blueberry fields. (There are records of the incidence of cancer in blueberry field workers.) At great cost of time, energy, and funds, the inquiry may be conducted with great competence over years so as to yield true knowledge of that frequency, and the quality of this knowledge may be entirely unaffected by how poorly selected was the posing of that frequency question as a means to answering the cancer question based on that frequency result.

For epistemological purposes, then, we must distinguish between two ways to assess the attempt to answer a whether question $\langle p? \rangle$. One is the assessment of that attempt from the point of view of the wider intellectual objectives that may be driving the sub-inquiry into that particular question, as when the attempt to answer the frequency question above is assessed for its potential to help with the cancer question. Second is the assessment of that attempt from the point of view of the likelihood of success in answering the specific question with success: that is, the likelihood of success in determining where the truth lies on that specific question. In the case of our frequency inquiry above, the attempt of the researchers rates very high in the second respect, but very low in the first respect. It rates high in execution and low in selection.

And isn't it obvious that knowledge involves an assessment just in the second respect? The first respect of assessment seems quite irrelevant to the theory of knowledge, even if it is extremely relevant to the theory of inquiry.

2.5. Unlike the basketball shot so poorly selected from across the whole court when there was no time pressure, many other athletic attempts involve both high risk and great creditable accomplishment. In basketball, there's the three-point score from near mid-court by the pure shooter under pressure from the clock and from defenders. In tennis there's a player running to make a nearly impossible return of a lob over their head at the net. Despite how high is the risk of failure, they properly run that risk given the circumstances and the great importance of that point.²

That player takes account of two sorts of risk. There is the salient risk of not returning the ball because of the extremely high probability that the topspin lob will elude them. And there is the background risk that the lights will go out, as they are extremely likely to do. The first sort of risk is taken into account, as the player knows that they are taking that risk, while proper tennis competence requires the player to run in any case. What about the lights? Isn't the player again just acting despite knowledge of a certain risk, where his competence requires that he dismiss the risk? Why say that the first risk concerns a salient condition whereas the second risk concerns a background condition?

Salient conditions are those that agents must feed into their cost-benefit analysis, as they calculate which of the options before them is to be opted for. That is how it is for the basketball shooter under time pressure near the end of a game. That is how it is for the tennis player who must scramble back despite how unlikely success may be given the heavy topspin on that low lob overhead.³

Is it like that for the condition of the lights and the role of this in the assessment of athletic performance? The condition of the lights *is* a condition that can enter into the conscious cost-benefit analysis of *whether to continue playing the sport on that occasion*. It is properly a part of cost-benefit analysis of whether to continue playing. And that decision can be properly assessed on the basis of the various factors relevant to it. However, the quality of that decision and of the players' continuing play may vary enormously, while that variation has zero bearing on the quality of the various domain-internal plays by the players.

2.6. Two further distinctions should be drawn as we turn our focus more fully to epistemology. First is one that derives from the bearing of the diachronic and the social for ordinary epistemology. It is these factors – the diachronic and the social – that explain the need for a threshold of competence above which we have competence *enough*. Enough for what? Enough, I submit, for storage. We properly do not store the evidential support enjoyed by our beliefs when they are initially acquired, nor any other features that may bear on such initial acquisition and its attendant degree of reliability. Soon after the belief's addition to our background knowledge, what we

²A better example, for those who know baseball, is that of the runner on third base who must run for home when an easy fly ball is hit to an awaiting fielder for an extremely probable easy out. The runner must run even if his chance of scoring is nearly nil, since the out is so easy. The point is that if the fielder somehow muffs the play, then the runner will score the winning run. So, the runner must run for home plate despite awareness that his risk of failure in that attempt to score is extremely high.

That runner takes account of two sorts of risk. There is the salient risk of not scoring because of the extremely high probability that the easy fly ball will be caught. And there is the background risk that the lights will go out, as they are extremely likely to do. The first sort of risk is taken into account, as the player knows that they are taking that risk, while proper baseball competence requires running in any case. What about the lights? Isn't the player again just acting despite knowledge of a certain risk, where his competence requires that he dismiss the risk? Why say that the first risk concerns a salient condition whereas the second risk concerns a background condition?

³And that is how it is for a baseball runner on third as the easy fly ball is hit right to the awaiting opposing fielder, even if here cost-benefit analysis yields the result that no change in the risk will affect the choice of whether to run for the score or not.

have is the belief itself, absent the factors that supported the initial acquisition. Such beliefs might later, over time, have bearing on a huge variety of potential decisions and actions, individual or collective. Others, and also our later selves, will need to presuppose that some minimal level of reliability will be attained by taking at face value the deliverances of memory and the testimony of others. This makes it proper for human communities to impose norms for belief storage, such that adherence to those norms would provide such reliability. And this works best if the requirement has the force of deontology and is not just one more factor subject to tradeoffs. So, the competence intrinsic to the epistemic domain is not just the competence to judge on normative tradeoffs generally, or to conduct cost-benefit analysis on all values that pertain in any way to the performance under consideration. Rather, in the epistemic domain we are to bracket non-epistemic considerations, and even consequentialist calculation pertaining to epistemic outcomes downstream from our judgment under consideration. The epistemic competence pertinent to knowledge is thus a very specific competence sensitive only to how likely our judgment (belief) is to be correct (true), given the competences that we rely on in that instance.

2.7. An additional distinction should be drawn as we turn to epistemology. Consider a baseball fielder's competence at a given time. This requires conditions of two sorts: salient conditions and background conditions. We already took note of these briefly; let us now consider them more expansively. This will lead us to a distinctive category of "ludic knowledge," and a corresponding "ludic epistemology."

Salient conditions include the ball's trajectory, and its spatial relation to the fielder, along with his tracking ability and foot speed. Such *salient* conditions must fall within certain parameters and the fielder must consider this at least quickly and implicitly. He would be negligent if he ignored it, and might be reckless for not giving it proper weight. This is how it matters that attempts be not just apt but fully apt, reflectively apt full well.

By contrast, lights known to be fragile can be properly ignored in assessing domain-internal credit for athletic success. Success remains about as creditable even when an athletic attempt is made in the knowledge of such risk. The condition of the lights is thus a *background* condition, one the agent can assume to hold good, even without knowing that it does, and no matter the degree of risk.

There is, however, a notable difference between domains that are ludic and those that are practical (and non-ludic).

2.8. Such practical domains concern the pursuit of non-ludic desiderata, including objective human goods, such as the production of food and of knowledge. Like domains that are ludic, practical domains allow default assumptions. The fact that a farmer failed to grasp the extremely high danger of a massive tornado or earthquake or tsunami need not reduce the credit earned for a great crop aptly produced through lengthy exercise of farming competence. A main point here again is that the quality of the farming performance can be *unaffected* by potential disasters that never materialize, no matter how probable a threat they may have posed.

Such practical domains still differ from the ludic in one important respect. A ludic performer's performance is not affected even if they *know* of the highly probable disasters. No increase in the known probability of a relevant disaster affects the quality of a ludic performance, so long as the agent *in fact* opts to ignore the possibility of such disasters.

Not so for agents in a domain aimed at securing some objective human desideratum, such as food, or knowledge. Consider such an agent heedless of some danger that they in

fact confront unawares. They may not be called upon to investigate the question of such a danger; they might just properly assume by default that the danger will not be realized. In this respect ludic and practical performers are alike. However, any practical performer who *knows* about the danger must take this information into account and cannot just continue to act within that domain on a default assumption that the danger will not be realized.

That is how it is for our farmers, and for other agents in pursuit of objective human goods, such as doctors, engineers, mechanics, accountants, and so on. And it now seems that way for the ordinary knower as well. Our ordinary default assumptions in epistemology differ importantly from their correlates in ludic domains. Default assumptions are less powerful in the epistemology of ordinary judgments and beliefs, and in practical domains generally. In ludic domains one can still assume by default that even *known* risks to success will not be realized. In practical domains, by contrast, such default assumptions can be proper *only if* one has no good reason to put them in doubt. No such default assumption is allowed once one knows of the pertinent risk and its severity.

2.9. Here we focus on the domain of ordinary, commonsense epistemology, along with domains of expertise, such as the legal, the medical, etc. But our telic virtue epistemology is quite naturally receptive to the admission of *ludic* epistemic domains. And with ludic epistemic assessment would come a kind of ambivalence in the assessment of whether one “knows” that one’s opponent is about to hit a tennis lob, for example, based on which one properly leans back towards one’s baseline. One surely has no *ordinary* knowledge of any such thing if one knows that the lights are poised to fail with near-certainty. And yet if the lights do *not* fail, then, along with the tennis quality of one’s leaning back and its aptness, there would surely be also a corresponding tennis-epistemic quality of one’s judgment that the opponent is about to lob, which could even amount to a sort of knowledge of that fact.

Along with the expert knowledge proper to the various domains of human expertise, we would now have various ludic domains of knowledge as well. And *these* ludic epistemic domains would admit not only epistemic default assumptions unaffected by lack of safety, but also epistemic default assumptions unaffected by the agents’ *ordinary* knowledge of such *lack* of safety. When I position myself based on knowledge of my opponent’s intention to lob, I can properly so judge if I can see the intention in their approaching stance, and in so properly judging I rely on a default assumption that the lights will not fail, *and* in making that judgment one might even rely properly on such a default assumption despite one’s *ordinary knowledge* of the dire condition of the lights.

2.10. How could we credit a tennis player with an apt angled winning volley without also crediting them with knowledge of the position and mobility of the opponent? But then, if we cannot deny them the aptness of that creditable winning shot despite the known likelihood that the lights will fail, neither can we deny them the aptness of creditable judgments about the opponent’s position, the judgments required for exercise of their competence in that winning shot despite the fragility of the lights.

2.11. The fact of such ludic varieties of knowledge is already a surprise, but the role of default assumptions in epistemology turns out to have additional surprising implications for virtue epistemology specifically, to which we turn next.

3. Grades of knowledge

3.1. Here is a familiar sort of case, which will soon raise some unfamiliar questions:

Simone is an ace fighter pilot with many years of experience. What she does not know is that she is now routinely subject to testing under simulation. While in the

dark about that, she now takes herself to sit in her genuine cockpit up aloft, seeing targets down below, and she is entirely right. But most often now she is only in a simulation cockpit that she cannot distinguish from its genuine twin. Can Simone really know to be true what she correctly and justifiedly takes to be true? Many answer emphatically in the negative.

Here true beliefs seem justified while short of knowledge. Most of us would take Simone to fall short that way, as supposedly does Barney (of fake barn lore). But there is also considerable dissent, so we may wish to take into account not only the majority's emphatic reaction but also the minority view. My own attempt distinguishes two sorts of knowledge: the animal versus the reflective. Simone can then be said to have animal knowledge without having reflective knowledge. We might thus hope to maximize explanation by having it both ways. But what exactly is the difference between these two sorts of knowledge?

Animal knowledge is belief that gets it right through competence rather than just luck. Reflective knowledge goes beyond that by requiring not only apt attainment of truth but also apt attainment of *aptness*. Reflective knowledge is belief whose aptness manifests the believer's *meta*-competence to attain aptness not just by luck but through competence.

Plausibly, Simone enjoys animal knowledge while falling short of reflective knowledge (full well). She has *merely* apt belief but lacks the higher-grade *full* aptness. Since she cannot tell whether she is in a genuine cockpit or only in the fake simulation, she seems to lack the competence to ensure that she would reliably enough get it right by trusting her perceptual input.

That turns out to be only a first approximation, however, as should become clear once we recognize more fully the role of default assumptions in epistemology.

3.2. If knowledge can be based on default assumptions, we can no longer so plausibly understand the case of Simone by distinguishing two forms of ordinary knowledge: by distinguishing animal from reflective knowledge. We can no longer argue that she has animal knowledge about the scene below *while lacking reflective knowledge*.

In order to attain reflective knowledge, all Simone needs is apt awareness that her beliefs about that scene below will be not just true but also apt. But what could deny her such knowledge?

Our default assumptions now come to the fore. Once they are admitted, we *cannot* after all explain how it is that Simone and Barney fall short of ordinary knowledge, at least not in the way we had supposed. We cannot just say that what they lack is reflective knowledge full well. Aided by default assumptions, they can attain reflective knowledge after all. Simone, for example, can no longer be said to lack knowledge that she is in the genuine cockpit, not if she has a right to default assume that if it appears to her that she is in a real cockpit, then that is where she is. What she still lacks is pertinent competence *secured* against simulation. Because of how easily she might be under simulation the competence provided to her by the genuine cockpit is not secure; she is not securely so located, since she might so easily be in the simulation cockpit instead.

That is why we will now appeal, beyond reflective knowledge, to *secure* knowledge full well.

4. Security explained

4.1. What is it for a belief to be thus "secure"? Is safe possession of the pertinent SSS competence required?⁴ No, one can surely know even when one's competence and

⁴My grateful thanks to John Greco for pressing this question.

even one's life are in great danger. A pilot could have perfectly secure knowledge even while in great danger of being knocked out or shot to death. The sort of safety that security requires is rather the following.

Secure knowledge

A given judgment constitutes secure knowledge only if the thinker is safe from the following fate: *losing* their pertinent complete SSS competence to so judge while at the same time *retaining* a disposition to make judgments when they "inquire" into that question *even absent any such competence* (on a broad notion of inquiry requiring only that the thinker take up that question, at least implicitly and representationally, if not through intentional conscious questioning or affirmation).

So, the key issue is whether too easily the thinker might have asked themselves the very same question and even delivered the same judgment in response, *without* manifesting any SSS competence (like the one manifest in the answer that they now actually give, and manifest also in the *success* of that judgment). This is a sort of safety lacked by Simone but retained by our knowledgeable pilot in dire danger of being knocked out or shot dead.

4.2. And we can now admit *grades* of knowledge. *Securely knowing full well* can be recognized as a higher grade than even *reflectively knowing full well*.

Although we rarely recognize and require that higher grade in ordinary epistemic assessments, it can still affect our intuitive responses to thought experiments, as it affects Descartes's intuitive response to his fourth skeptical scenario, where we are creatures of a lesser creator, or of a random sequence of causes. This seems thus a higher knowledge involved in our intuitive *denial* that Simone can "really know," as she might so easily have been in the simulation cockpit rather than where she is up aloft. The higher knowledge then lacked by Simone is one that requires security, beyond aptness full well.

Take then any remaining intuition according to which the threat of simulation blocks her from *really* knowing as she views the scene below her cockpit. How is any such remaining intuition to be accommodated? No longer by denying that Simone reflectively knows full well. Perhaps rather by recognizing that what she lacks is *secure* knowledge, even when her judgment is apt (her alethic affirmation *fully* apt), so that she knows reflectively.

5. A closer look

5.1. What now distinguishes the following two Simones? And let's suppose them both to be veteran pilots whose training decades ago included simulation, but always knowingly, and never repeated until now.

First is a Simone⁻, who is *now* unbeknownst to her more likely to be in a simulation than in a real cockpit, as unbeknownst to her she is now regularly tested under simulation.

Second is a Simone⁺, who is even now in no danger of being under simulation, having been a fully functioning fighter pilot for many years, completely safe from simulation, as she still is now, when for one thing she is not remotely a candidate for such testing.

If we are still pulled intuitively to think that Simone⁺ is epistemically better off than Simone⁻, if we still think that Simone⁻ falls short and does not know, or does not

know as well as does Simone+, it is tempting to explain that difference through the difference in security. The two Simones differ in that Simone– might far more easily be misled by simulation, which lowers the epistemic quality of her pertinent beliefs (even if she suspects no epistemic danger, and has no reason for suspicion).

If a belief is to constitute *secure* knowledge, more is required than just that it be a belief whose correctness (truth) manifests the thinker's competence, while its aptness does so as well. Security goes beyond aptness and even beyond reflective knowledge full well.

5.2. But is this too facile?⁵

It does seem facile to just declare Simone to lack a sort of “knowledge,” namely “secure” knowledge full well. After all, we could with equal facility declare that there is a sort of “sensitive” knowledge lacked when one believes that one's lottery ticket will lose.

And similarly for a “safe” knowledge that *p*, one requiring, of a true belief that *p*, just that not easily would one then believe that *p* without it being so that *p*.

And we could even propose, on a similar basis, a notion of JTB knowledge. This is a distinctive sort of knowledge that one has of the fact that *p*, when one believes that *p* with both truth and justification. There was a lot of intuitive support for this until its ostensible refutation by Edmund Gettier. Why not say here again that the best way to account for that intuitive support is to grant that there is a distinctive sort of knowledge, JTB knowledge, and that supporters of JTB were just responding to that?

Such facile methodology explains intuitions, and the clash of intuitions, by alleging that intuiters just focus on different real varieties of knowledge, and then voice their intuitions in ways understandable through linguistic explanations.

If these latter moves seem so facile for supposed special varieties of “sensitive” or “safe” knowledge, or even JTB knowledge, why is it less facile to propose a sort of “secure” knowledge full well?

6. What unifies the hierarchy?

6.1. Does security belong with aptness and full aptness among the metaphysical categories pertinent to epistemology? *More generally*, does security belong with aptness and full aptness among the metaphysical categories pertinent to the assessment of human performance in the many domains where humans perform telically? At the core of such assessment are questions of credit and blame, and of responsibility even of the sort that involves only *attributability* and not necessarily *accountability*.

6.2. Security has been given a place in our metaphysics of epistemology, so as to explain why it is that Simone would not really know if she might so easily have been under simulation. What she would lack is a demanding sort of knowledge, *secure* knowledge full well.

But wait. Wasn't the domain of ordinary judgment and belief supposedly sealed off from external dangers in ways familiar from human performance domains generally, be they athletic, artistic, professional, etc.? How can we now be told that the dimension of security *does* affect epistemic quality of domain-internal performance, that in epistemology a belief is somehow epistemically more creditable when it is not only fully apt but also secure?

Security requires not only the *actuality* of pertinent background conditions but also their *safety*. And to say that a domain is sealed off from external dangers is to say that

⁵My grateful thanks to Chris Kelp for pressing this question.

agents can perform *flawlessly* in that domain, based on a default assumption that the pertinent background conditions are satisfied, even when this is *far from being known* and even *far from being safe*. That's how it is for the baseball fielder with respect to the fragility of the lights.

But now we're told that it's different for the pilot Simone, as concerns the fragility of her situation. The epistemic quality of the perceptual judgments made by Simone when aloft is now supposedly affected by their security, and not just by their aptness or full aptness. And isn't this to say that the domain of ordinary epistemology is not sealed off after all? Isn't it to say that the safety of Simone's situation does after all affect the epistemic quality of her domain-internal ordinary judgments and beliefs?

6.3. What to say? First of all: Good point! This requires attention; and here now is a start. Consider first *domain-defining* conditions: that is, any condition constitutive of performing in a given domain, in such a way that performing in that domain requires satisfying that condition. Thus, a condition for playing a given sport at night may require that one play in a lighted field. Once the lights go out and the field is *too* dark, participants are no longer in conditions where they can really play that sport. The domain of baseball, for example, requires a lighted field, one *well-enough* lit, either through natural daytime light or through artificial lighting.

When agents in a domain are ready for performance in it, they must assume that the conditions for so performing obtain. Even when they do not check to make sure, they can still assume by default that those conditions do hold; they *can* assume this by default in the following sense. If they do assume it by default and the conditions do hold (no matter how unsafely) then the quality of their domain-internal performances is not at all affected by any lack of safety in those conditions. Nor is the quality of their performance affected by any lack of knowledge on their part, about the holding of those conditions.

Up to now, by hypothesis, our Simone— has been well into a long career as an ace fighter pilot. Simulations have been entirely in her remote past as a pilot trainee. Nevertheless, if now pilots are subjected to extensive and surreptitious simulation testing, her SSS competence becomes insecure. In particular her situational S becomes highly insecure. But might she not *assume by default* that her competence is secure? Why would she be subject to a charge of negligence or recklessness, if athletes in a night game with fragile lights escape any such charge?

Yet the quality of epistemic performance by Simone— does very plausibly seem to suffer. Intuitively she does not know as well as when she is not in such great danger of simulation. So, the pertinent epistemic domain seems not sealed off after all. It is not sealed off from the influence of insecurity on the quality of domain-internal performance.

Let us have a closer look.

6.4. Let us compare Fielder (a baseball fielder in a night game, under fragile lights) and Simone (the veteran fighter pilot who might now too easily be under simulation, and let us focus now more specifically on the case of Simone—, who is now more often in the simulation rather than in the genuine cockpit). What threatens Fielder's competence threatens also his ability to perform in a proper baseball domain. His performance marred by darkness would no longer belong in a proper baseball domain. By contrast, whether in the simulation or up aloft, Simone *would* still issue judgments, still make affirmations aimed at aptness, even if they would be defective. Even under surreptitious simulation, she would of course still make genuine judgments.

Distinguishing that way between Fielder and Simone is in line with our account of secure knowledge. Now the *ordinary* epistemic domain is not after all sealed off from

failures of security, in the way of ludic domains. In the case of epistemology, agents can fail a requirement of the sort of security explained above in our account labelled “Secure Knowledge.” Thinkers can continue to make judgments even having lost their pertinent epistemic SSS competence. And, in respect of epistemic quality, insecure judgments fall short of those that are secure.

By contrast, the baseball fielder’s attempted catch is not downgraded because of the loss of security consequent on the fragility of the lights, nor does it even matter if the fielder knows of the fragility. Fielder and Simone differ, moreover, in a way that helps explain that distinction: Fielder loses the ability to engage in his domain when the lights go out, but Simone does not lose her corresponding ability when her danger materializes and she is under simulation. When the lights go out, Fielder can no longer engage in the baseball practice through attempted catches, having lost the preconditions for a proper game, as does a chess player whose chess board is whisked away. By contrast, Simone can still engage in epistemic practice through genuine judgments, even when she is under simulation.

6.5. What of familiar skeptical scenarios like the evil demon, the brain in a vat, or the Matrix. These are radical by comparison with Simone’s simulation, which raises the question of whether the victims retain the capacity to make genuine judgments. My own preference is to answer in the negative: radical victims are unable to make genuine judgments. To me this seems plausibly connected to the metaphysical detachment from society that comes along with their detachment from their surrounding world. Human beings and human knowledge plausibly require the metaphysical connection with human community and an external world that radical scenarios would tear away from them.

Once someone is envatted, we can speak of their doings in the past tense but no longer in the present tense (not strictly). If we say that they are still there, that’s like saying that someone is being buried, at “their” burial, or that “they” did certain things in their life. Those can be proper and true things to say, while having no real existential import, in the sense of referring to someone still in existence. By hypothesis, there *was* someone so-and-so, who *did* various things, but strictly no-one still in existence did those things.

In the case of a live brain in a vat of nutrients, a BIV, there is still a brain in existence and it was formerly someone’s brain, when there was someone in existence whose brain this was. But now no-one has that brain, and there’s no-one who has mental properties through the doings of that brain. The brain has various properties, but none is or determines a property of a person, since there is no longer any person whose brain that is, so that they can be determined to have any properties through the properties of the brain.

Why not? Well, in part because there’s no entity that owns the brain while playing the social and causal roles required for the constitution of a person. Advanced technology might give such a detached brain the chance eventually to underlie the existence of a person. Maybe technology would provide new ways for a brain to relate to a wider world so as to play the roles of intercommunication and collective action proper to a fully fledged person. And that might even count as the resurrection of the dead. But until that happens, there is no longer any person constituted by such a BIV, nor are there any doings of a person that are constituted by the doings of that BIV.

But why not? Why not rather say that the person remains so long as there is an entity that retains a unified complex set of *potentials* whose realization would ensue upon its being properly “reconnected” by surgery or by some more advanced technology. All we would need for the constitution of a person is that there be such an entity.

Yes, we *could* think that way. But consider the collection of atoms that constitutes that brain. Is it not true of that collection as well that it has such potentials, even

when the atoms are scattered. All it would take is the construction of that brain by bringing about the proper interrelation of those atoms, etc.

Suppose our intuitive response is that the set of atoms is just too remote from a full human being and person. Consider its potential to gain the proper connections with the external world and with external communities (actual or potential) that would enable it to acquire the causal and social roles required for the constitution of a real person. We are told that this potential is too remote for the collection of atoms but not too remote for the BIV.

Is that supposed to be obvious? What then is the difference that makes so crucial the supposedly greater remoteness of the scattered collection of atoms, its greater remoteness from full personhood, by comparison with the lesser remoteness of the coherent BIV? Something must account for why the BIV *can* and does but the collection *cannot* constitute a person, not while scattered. There is supposedly a difference in the sorts of potentials retained by the brain once detached from the human body and the sorts of potentials retained by the collection once its member atoms are scattered. But we've been offered no clue as to what sort of difference that could be.

Is it supposed that we can directly intuit that the BIV *would* metaphysically underlie a person, whereas the collection of atoms would not? Why couldn't the collection underlie a person in a condition analogous to being totally asleep or otherwise unconscious?

Why couldn't the collection of atoms simply no longer underlie a person, no longer constitute a human being or a person, nor even a thinker, not an actual thinker of thoughts nor even a potential one? And once we say that of the collection, why not say it of the BIV as well?

Compatibly, we could grant that there are things quasi-mental that go on in the BIV and in any higher-order being that might be constituted by the BIV. That might be like saying that there are quasi-assertoric things that go on in the performance of a play. There are many complex sayings complexly arrayed with other sayings by other players, with all the players in complex sequential interaction. Still there may be not a single assertion in the whole two-hour performance.

Similarly, there might be lots of quasi-thoughts by the BIV (or by any higher-order entity constituted by that BIV), without a single real assertoric thought in the whole lot. But why? Well, first: why not? Is metaphysical truth supposed to be open to the bare eye of the mind just straightforwardly as is the presence of hands to competent visual perception?

And what do we lose, really, if we take another path, if we say that to us it seems no less obvious, or no more unobvious, that a BIV is no real thinker? What if we find it plausible that real thinking requires proper insertion into a larger social world involving insertion in an external world with channels through which we gain access not only to hands and fires but also to other minds, for interaction and collective action?

What do we lose, really, if we deny that a BIV can be a proper, real thinker? Well, one thing we seem to lose is radical skeptical scenarios, such as that of the BIV, and, by similar reasoning, others as well, such as that of the Evil Demon, and that of the Matrix. But are these things we really want? Or is their loss part of the therapy that might give us proper intellectual peace?⁶

⁶This response to radical skepticism is notably similar to the response to dream skepticism in "Dreams and Philosophy," my Presidential Address to the APA, which also figures as the first chapter of *A Virtue Epistemology* (Sosa 2007). Both responses dispute the claim of the skeptic that in their adduced scenario we would be misled into false judgments. What is disputed is that in the respective scenarios we would so much as make judgments. However, neither of these is a universal all-purpose response to all forms of skepticism. My *Reflective Knowledge* (Sosa 2009) has broader scope.

6.6. All that being so, should we conclude that the commonsense, ordinary epistemic domain is a ludic domain? Is that how we can properly ignore radical skeptical scenarios? Is our default dismissal of such scenarios quite like the fielder's dismissal of the possibility that the lights might fail? True, these dismissals *are* similar, and have been argued to have a similar basis. The basis in each case is that if the truth that we assume by default to be true were false instead, that would deprive us of the ability to perform in the relevant domain. Thus, the failure of the lights in the night game would block us from being able to make baseball plays, such as attempted catches. Similarly, our being envatted, or Matrixed, or victimized by an Evil Demon, would block us from making proper judgments.

If so, the ordinary epistemic domain is then strikingly similar to ludic domains, as we have seen. But the respects of similarity do not suffice to entail that the ordinary epistemic domain *is* itself just a ludic domain. Epistemic domains *can* indeed be ludic, as is the epistemic subdomain of crossword puzzles. But when the successful pursuit of truth amounts to knowledge there is no apparent reason why we should always consider that to be success in a game, even if there are shared elements. There is no apparent reason to think that the pursuit of truth and its apt attainment must always amount to the playing of a game, any more than to think that the successful production of a good crop is success in a game.

6.7. Consider categories pertinent to degree of attributability/responsibility, and opposed to metaphysical luck. In this regard, the category of security (of secure fully apt attempt) stands in stark contrast to any alleged categories of sensitive attempt or safe attempt.

Sensitivity and safety, as applied to attempts, are more general versions of sensitivity and safety of beliefs.

A sensitive attempt is *one that the agent would not have made if it would have failed*. But this does not properly connect the success of the attempt to the agent as a source of that success. After all, compatibly with the sensitivity of their attempt, perhaps they would have omitted that attempt *regardless*, i.e., even if it would have succeeded. Why they *would have omitted* that attempt regardless is perhaps because they are *unable* to make *any* such attempt, disabled as they are by some injury.

Something similar applies to safe attempts. A safe attempt is *one that the agent would have made only if it would have succeeded*. But this again fails to properly connect the success of the attempt to the agent as a source of that success. After all, it may be that the success would have come forth *regardless*, i.e., that the success was bound to happen, and that the source of its happening had nothing to do with the agent's effort. Thus, the agent may attempt to stay awake throughout a lecture, and may do so safely, in that he would have made that attempt only if it would have succeeded: that is, only if he would have stayed awake throughout the lecture. However, what ensured that he would stay awake had nothing to do with his effort. Rather, it was due just to the highly caffeinated cup of coffee that he drank (while believing it to be decaffeinated), where a huge conspiracy has arranged several layers of backup protection to keep him awake for that important lecture.

Nor is either of *sensitivity* or *safety* plausibly constitutive of any of the categories on the higher levels of our hierarchy, the way *success* and *competence* are so constitutive.

All of that being so, neither the sensitivity nor the safety of our attempts plausibly belongs in a metaphysical hierarchy designed to reveal levels or degrees of attributability/responsibility, or (in the opposite order) levels of telically pertinent *luck*.

Secure knowledge does seem properly to occupy a higher level of our epistemic/metaphysical hierarchy. Secure knowledge (full well) rules out additional luck that is

plausibly credit-reducing (credit-reducing in the sense of attributability-reducing, with no further axiological or deontic implication, so that a perfect murder could still be fully “creditable” to the agent).

6.8. That is still compatible, however, with our frequent focus in ordinary human affairs on levels of attributability that do not require security. Despite this focus and how it affects ordinary epistemic and other performance assessments, a clear failure of security in a thought experiment can still bear on our intuitive responses to the question of whether the protagonist *really* knows, whether they are fully responsible, whether their action is fully attributable to them. An outcome can be sufficiently attributable for all practical purposes of ordinary human life, even if it is not sufficiently attributable for purposes of full Cartesian assessment. Similarly, a box might be empty for all practical human purposes, even if it might still be full of air.

Moreover, a Cartesian-light attributability/responsibility requires, not the *radical Cartesian* security of one’s pertinent SSS competence, but only the sort of ordinary security that one has when not easily would one have lacked one’s pertinent triple-S competence. Even if we do not share an interest in full Cartesian assessment, we can still recognize that it is when our Simone fails even the Cartesian-light requirement that it is so plausible that she fails to *really* know, by failing to earn full credit for her athletic successes.

And yet ... Put yourself in the cockpits of nearby flyers in formation next to Simone on a genuine mission. The maneuvers that they make as they fly their dangerous mission over enemy territory will require constant fast assessments of Simone’s state of mind and what she is likely to do. How would they conceptualize this if not by advertizing to what she sees and hears, what she knows accordingly, and how they can trust her to perform in the light of that assessment of her mind. Within the sub-domain of their mission, they make default assumptions like the assumptions made by a tennis player who “knows” what an opponent is likely to do (despite the acknowledged extreme fragility of the lights). The knowledge that her fellow pilots attribute to Simone during their mission is of course no mere “ludic” knowledge. They are not just playing a game or engaged in a sport. Their mission is deadly serious. Yet there seems a close analogy between the ludic knowledge of a tennis player about the opponent’s actions and the “mission-relative” knowledge of the fellow pilots. In both cases there are domain-defining default assumptions relative-to-which a kind of domain-internal knowledge is properly possessed and attributed even despite its clash with domain-external *ignorance* on the part of the same agent (Simone as it might be) about the very same propositions. Simone does not have *secure* knowledge on which to base her maneuvers and her fellow pilots on the mission need not attribute such knowledge to her as the mission unfolds successfully and creditably. It suffices that they attribute to her knowledge on a lower level (the animal or the reflective) based on which attributions they can successfully conduct their mission.

6.9. In sum, what distinguishes belief that rises to the level of knowledge in our epistemic/metaphysical hierarchy is that it is belief whose *success* is significantly *creditable* (*attributable*) to the believer, one for which the believer is attributably responsible as their own doing, and not just something that comes about by (agent-external) *luck*. However, there are distinct ways of attaining such creditable success, ways that form a metaphysical hierarchy. And what *orders* that hierarchy? What orders the hierarchy, I submit, is the degree to which the success is creditable to the agent.

At the lowest level is found the *merely* apt athletic affirmation, whose success (truth) is due to the first-order competence of the thinker. That is the level of animal knowledge. (This can be a *sub-credal* sort of “knowledge,” as is the knowledge of eye-exam

subjects who turn out to get it right with supreme reliability even when they “guess” at the lowest reaches of their eye-chart, where the letters are extremely small.)

Immediately above that lowest level is the *fully* apt alethic affirmation, the apt *judgment*, whose aptness is due to the meta-competence of that thinker, the meta-competence that enables them to discern when their alethic affirmations would be not just true but also apt. This is the level of *reflective knowledge full well*. Here we have a significant increment of credit to the thinker, one aligned with a reduction of the pertinent luck in the thinker’s attainment of success.

Finally, we have found a third level of knowledge, above reflective knowledge full well. Here the thinker does attain reflective knowledge full well, and *in addition* this attainment derives from competences retained safely, not just through luck. This is the level of *secure knowledge*.

What unifies these three levels of our hierarchy, as distinct levels of *human knowledge*, is that on these levels the thinker attains an epistemic success (truth, or aptness) attributable to them, as really their own doing. Coordinately, such success corresponds decreasingly to adventitious external luck. This holds good all along the ascent of attitudes from the animal level of the merely apt, to the reflective-full-well level of the fully apt, to the *securely* reflective-full-well level attained through competences retained safely and not just by luck: i.e., competences that would not too easily have been missing.⁷

References

- Sosa E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, Vol. 1. Oxford: Oxford University Press.
- Sosa E. (2009). *Reflective Knowledge: Apt Belief and Reflective Knowledge*, Vol. 2. Oxford: Oxford University Press.

Ernest Sosa is Board of Governors Professor at Rutgers University. He has published papers and books in epistemology, including most recently *Judgment and Agency* (Oxford University Press, 2015) and *Epistemology* (Princeton University Press, 2017).

⁷In the course of writing this paper I have discussed the issues in it and related issues with John Greco, Chris Kelp, Matt McGrath, Ram Neta, and David Sosa, to all of whom I am very grateful.