## Preface

This Element defends a 'stratified' version of virtue epistemology. Virtue epistemology is built around a simple idea: knowing a fact essentially involves believing it truly through the exercise of ability; knowledge is thus always and everywhere an *achievement* on the part of the knower. This simple idea is theoretically powerful. Section 1 shows that, working with just this simple 'knowledge = truth through ability' slogan, we can get *far* better results in epistemology – demonstrated across a spectrum of theoretical test points and cases – than critics of virtue epistemology have appreciated thus far. And it is shown to do better than notable competitor proposals, such as safety-based accounts of knowledge. But as we'll see, there is, at the end of the day, only so far you can get with just one *level* of knowledge. Some residual problems remain, no matter what moves one tries to make while working with a single, 'uni-level' virtue epistemology.

Against this background, Section 2 follows Ernest Sosa's lead and considers what we can achieve by adding a second 'level' to our virtue-theoretic picture – a distinction between *animal* and *reflective* knowledge (roughly: between knowing, and knowing that you know – viz., knowing *knowledgeably*). It is shown how a multi-tiered, 'stratified' version of virtue epistemology gets better results, all things considered, than a more traditional, 'uni-level' virtue epistemology, and with fewer theoretical costs. In the course of developing this multi-level idea, Section 2 takes us beyond Sosa and offers a new account of reflective knowledge – carving out a place for both *descriptive* and *predictive* reflective knowledge, and the distinct theoretical roles each plays.

Section 3 then shows how the account – which by this point recognises a stratified picture of *knowledge* – is improved further yet through the introduction of *stratified beliefs* into the picture; on this view, (put roughly) some beliefs *constitutively aim* higher, epistemically speaking, than others. Working with this idea, Section 3 then develops a new account of the highest grade of knowledge – fully apt judgement – which (motivated along the way by several critiques of Sosa's account) incorporates theoretical innovations concerning both (i) *level-connecting* (between the animal and reflective levels), and (ii) how it is that 'high grade' knowledge interfaces with epistemic risk and background beliefs that we can non-negligently take for granted.

The resulting picture is a new and improved version of virtue epistemology, one that is situated broadly within the Sosan tradition, but which takes us beyond it in important new ways. I'm grateful to Stephen Hetherington – series editor for Cambridge University Press's *Epistemology Elements* series – for encouraging me to write this Element for the series, and to two helpful

anonymous referees at Cambridge University Press and to Stephen – for helpful feedback. This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 948356, KnowledgeLab project, PI: Mona Simion). Thanks also to the Leverhulme Trust (RPG-2019–302) and the AHRC (AH/W005077/1) and (AH/W008424/1) for supporting this research.

## 1 Virtue Epistemology: One Level Is Good

The core knowledge thesis embraced by virtue epistemologists is that that propositional knowledge is type-identical with *apt* belief – viz., belief whose correctness is because of ability. The devil is in the details of the view (which we'll jump right into in this section), but the basic idea is just as simple as it sounds: some of your true beliefs are due to luck; those aren't knowledge. Those that are due to ability (and only those) are the ones you know. That's the crux of the idea. Can a theory of knowledge so simple be extensionally adequate?

If you've done a bit of epistemology already, you might think not. There are some well-known lines of argument in the literature that hold that apt belief is neither necessary nor sufficient for knowing.[1] Moreover, there are sceptics about the very project of analysing knowledge.[2] But we'll see that the situation turns out to be much better than critics have appreciated. This section is going to show just how well the simple idea that knowledge is apt belief can do, without any extra bells, whistles, or (as we'll add in later sections) 'levels'.

The plan for this opening section will be to begin by showing how the knowledge = apt belief equivalence allies itself naturally with two related theses to form a kind of 'core triad'. The core triad, we'll see, holds up really well against the competition once we test its explanatory power across a relatively wide testing ground of cases and problems. Along the way we'll distinguish two substantive ways of characterising the template idea that knowledge is apt belief – due to John Greco and Ernest Sosa, respectively. The idea that knowledge = Greco-aptness does well, but we'll see the idea that knowledge = Sosa-aptness does even better (and both do better in our testing ground than notable competition, including Pritchard's anti-luck virtue epistemology). Methodologically, we'll keep a kind of running scoreboard throughout the section, with a final scoreboard at the end. The final scoreboard – while it shows just how well the 'knowledge = Sosa-aptness' view performs – also leaves us with some lingering questions, which will set the scene for the next section, where we begin to see advantages of a multi-tiered account of knowledge.

[1] See, for example, Lackey 2007; Pritchard 2012. [2] See the Appendix to Section 1.

## 1.1 The Basic Core Triad

The very idea that knowledge is type-identical with apt belief already – and before we get into different substantive glosses – commits its defenders to two closely related theses. First, consider that *aptness* – a property a belief has when its success is because of ability – is a *normative* assessment: by calling a belief apt we evaluate the belief relative to an implicit standard governing the kind of attempt a belief is. From the core idea that knowledge is type-identical with apt belief, we are tacitly signed on to the thesis that knowledge is a *normative* kind, as opposed to (say) a natural kind.[3] Second, when any aim (an archery shot, a dance performance, etc.) is secured not just luckily but through skill or ability, the success is *thereby* an *achievement*, where (pre-theoretically at least) we take the (attributive) goodness of an achievement to outstrip the goodness of the mere success. Qua achievement, then, knowledge is not merely normative, but it also has some (defeasible[4]) normative 'oomph'.

With these ideas in play, we can now see that the core idea that knowledge is apt belief (viz., the virtue epistemologists's core knowledge thesis (CKT)) is best understood as the 'core' of a key triad of claims, all three of which are capable of doing explanatory work.

*Core Triad (Virtue Epistemology)*

- **Core knowledge thesis (CKT)**: Propositional knowledge is apt belief.
- **Normative kind thesis (NKT)**: Knowledge is a normative kind.
- **Cognitive achievement thesis (CAT)**: Knowledge is a (species of) cognitive achievement.

This package is often, *by default*, endorsed by those who accept also the orthodox 'uni-level' thesis about grades or levels of knowledge:

- **Uni-level thesis**: There is one and only one grade of propositional knowledge.

For the virtue epistemologist, then, the uni-level thesis implies a commitment to thinking that knowledge-qua-apt belief is a *single normative kind*, and that the achievement one attains when knowing is always and everywhere just the achievement that is associated with the normative kind of *apt belief*.

[3] Normative kinds, unlike natural kinds, are type-individuated in an irreducibly normative way. Social kinds might be either normative or natural kinds, though there is disagreement on this point (Bird and Tobin 2022, section 2.4).

[4] An achievement's being trivial or evil might implicate that it is all-things-considered of little (or bad) worth. See, for example, Carter 2023; Sosa 2021, chapter 2.

But what exactly does apt belief involve, in virtue of requiring the belief be a kind of 'success from ability'? Let's now think about one popular answer to this question.

## 1.2 Knowledge as Greco-Aptness

One prominent defence of the core triad – paired with the uni-level thesis – is due to John Greco (2010), whose book *Achieving Knowledge* is centred around the simple slogan that knowledge is always and everywhere *success from ability.*

To put some substantive meat on the bones, we need to know precisely what Greco means by both 'ability' and by 'attributable to'. He offers his own distinctive account of both. *Cognitive abilities*, for Greco, are *environment relative* stable dispositions to believe truly reliably. For example, you might right now have a visual-perceptual ability that you exercise to correctly ascertain the colour of the wall in the room, but for Greco you wouldn't possess or exercise this ability if you entered a house of illusions, where visual perception is unreliable.

Regarding *attributability*: For Greco, a belief's correctness is *attributable* to ability when ability (rather than, for example, luck) is the *most salient part of a causal explanation for why the subject believed truly.*[5] Greco's particular way of defending CKT, then, involves the following substantive view of what aptness involves; for convenience, call this *Greco-aptness*:

> **Greco-aptness** A subject's *S*'s belief that *p* is apt iff the most salient part of a causal explanation for why *S*'s belief that *p* is true is *S*'s (environment-relative) stable disposition to believe truly.

By identifying knowledge with Greco-aptness, Greco holds that the kind of *achievement* one has when knowing is one that requires their environment-relative abilities to most saliently (alternatively: primarily) causally explain why their belief is correct. If something else (luck, a special helper, etc.) is comparatively more salient as part of a causal explanation for why they got it right, or if one's environment-relative abilities partially but don't *primarily* explain why they got it right, then their success is not a cognitive achievement; their belief is not apt, and they fail to know.

[5] This is a standard simplification of Greco's 2003–10 view, which we find in, for example, works by Pritchard (2012) and Lackey (2007). One reason for opting for this simplification in presentation is that Greco takes the mechanisms governing salience to support picking '*one* partial cause rather than another' which is 'important' within a wider causal explanation (see Greco 2010, 74). Greco's position takes a different shape in another work (Greco 2020b). Given the attention the earlier view has received, combined with limited space, this Element will focus on the 2003–10 view.

## 1.3 Greco-Aptness and Gettier

Perhaps the most impressive advertisement for Greco's identification of knowledge with Greco-aptness is the way his view easily 'rules out' standard Gettier cases as cases of knowledge. To see how this works, just consider the following simple Gettier-style case:

> SHEEP IN THE FIELD: Roddy is a farmer. One day he is looking into a field near-by and clearly sees something that looks just like a sheep. Consequently he forms a belief that there is a sheep in the field. Moreover, this belief is true in that there is a sheep in the field in question. However, what Roddy is looking at is not a sheep, but rather a big hairy dog that looks just like a sheep and which is obscuring from view the sheep standing just behind.[6]

In SHEEP IN THE FIELD, Roddy has a justified true belief that *there is a sheep in the field*. But – as this case bears a classic Gettier structure – Roddy doesn't *know* there is a sheep in the field. Why not?

From Greco's perspective, the answer is simple: in SHEEP IN THE FIELD (as well as other Gettier cases), the subject, *S*, believes from an ability and has a true belief, but the fact that *S* believes from an ability is not the most salient part of a causal explanation for why *S* has a true belief – thus no Greco-aptness and a fortiori no knowledge. This is so even though *S*'s believing from ability is *a part* of the total set of causal factors that give rise to their believing truly.

A critic might press back: Why *isn't* cognitive ability the most salient part of the causal explanation for the subject's getting it right in Gettier cases? What are the mechanisms governing explanatory salience that would get this result? According to Greco (2008, 2010) explanatory salience is partially a function of our interests and purposes, and therefore, a function of what is normal or usual in light of these interests and purposes. Given our interests and purposes as information-sharing beings (viz., we as a default want to share and receive good information), our intellectual abilities have a *default salience* in explanations of our true belief. However, as the thought goes, in Gettier cases, this default salience is trumped by something abnormal in the way that the subject acquires a true belief. In effect, Gettier cases involve something akin to a deviant causal chain.[7]

So far, uni-level virtue epistemology's core knowledge thesis is looking good. It not only deals with Gettier cases, but does so in a straightforward way. As far as traditional Gettier cases go, the view gets full marks.

[6] This is Pritchard's (2009) variation on a Gettier-style case due originally to Chisholm (1977).
[7] For criticism, see Pritchard (2008).

## 1.4 Value of Knowledge

One of the trickiest contemporary problems in the theory of knowledge concerns the relationship between the nature of knowledge and the *value* of knowledge. No one denies that knowledge is valuable. But why is it valuable? A knee-jerk answer here holds: because it can get us what we want! Put another way, knowledge is *instrumentally* practically valuable. True, however, we *also* think that knowledge is more valuable than *mere* true belief[8] – viz., true belief that falls short of knowledge. So a sharper question is: *in virtue of what is knowledge more valuable than mere true belief*?

Here is where things quickly complicate. A lesson from Plato's *Meno* is that mere true belief will get us what we want just as well as knowledge. After all, one who truly believes that a given road leads to Larissa is as well served as one who knows that it does.[9]

Explaining why knowledge is more valuable than mere true belief is especially difficult for reliabilist accounts of knowledge, which identify knowledge with reliably produced true belief.[10] This is because reliabilists conceive the difference between knowledge and true belief that falls short of knowledge as a difference in the reliability of the source. But the reliability of a source, as Zagzebski (2003) and Kvanvig (2003) have argued, cannot add value to its product.[11] The value of a good cup of espresso is not increased by the fact that it was made by a reliable espresso machine. (Just consider: a cup of espresso with the same intrinsic qualities, but made by an unreliable machine, would have exactly the same value.) The conclusion, then, seems to be that reliabilism cannot explain why the value of knowledge exceeds that of mere true belief; if (as we think it is) knowledge is more valuable than mere true belief, its value must be explained in some other way.

At this point, a uni-level virtue epistemologist such as Greco has a promising card to play. Remember that part of the virtue epistemologist's 'core triad' of claims includes the *cognitive achievement thesis*:

**Cognitive achievement thesis (CAT):** Knowledge is a (species of) cognitive achievement.

Cognitive achievement thesis now comes in very handy. If knowledge is always and everywhere an *achievement*, then we can make sense (easily, in fact!) of

[8] The core idea needn't implicate that we think knowledge is significantly more valuable than a corresponding mere true belief in the same proposition. For discussion on different ways to capture the driving intuition, see Greco 2010, chapter 6; compare, Hetherington 2018.
[9] Compare, however, Goldberg 2023; Olsson 2007; Williamson 2000.
[10] See, for example, Goldman 1999. [11] For criticism, see Carter and Jarvis 2012.

why it is better than *mere* true belief. All we need to get that result is to pair CAT with the independently plausible *value of achievements* thesis:

**Value of achievements thesis (VOA):** All achievements – successes from ability – are finally valuable.

'Final' value is *non-instrumental* value – viz., value something has not just for the sake of something else.[12] The idea captured by VOA – viz., that achievements, as such, have final (non-instrumental) value – is easily motivated by reductio. Consider that if achievements were *merely* instrumentally valuable relative to the target success, then we'd have no reason to prefer, for example, an archery shot that succeeds through skill to one that succeeds just through luck. But we do! And insofar as we do, the best explanation here is that an achievement (a success from ability) is not valuable only insofar as the relevant success is valuable. From CAT, along with VOA, we derive (with no other premises needed) the conclusion that knowledge is finally valuable, and a fortiori, more valuable than *mere* true belief.

At this point, uni-level virtue epistemology is looking great: full-points for *both* the Gettier problem and the value of knowledge problem.

## 1.5 Temp-Style Cases (Safety without Aptness)

Uni-level virtue epistemology has got some momentum. Let's keep riding it – right up until the point where we see a good reason not to. Recall that one of the uni-level virtue epistemologist's credentials is that it offers an elegant way to rule-out knowledge in Gettier cases. Well, so does another competitor type of view, one that appeals to *safety* rather than aptness to do the trick.

**Safety condition (SC):** *S*'s belief is safe if and only if in most nearby possible worlds in which *S* continues to form their belief about the target proposition in the same way as in the actual world, the belief continues to be true.[13]

In Gettier cases like SHEEP IN THE FIELD, Roddy's belief is not apt, but it's *also* not safe, with reference to SC. Very easily, Roddy believes falsely in nearby worlds (that there is a sheep in the field) when we hold fixed the way he formed the target belief in the actual world. So, the thesis that knowledge requires safety seems to do just as well as uni-level virtue epistemology does in ruling out Gettier cases as cases of knowledge.

[12] See, for example, Rabinowicz and Rønnow-Rasmussen 2000; and, for a classic presentation of this idea, Aristotle in the *Nicomachean Ethics* I.7.
[13] See Pritchard 2005.

Now *this* observation might lead one to reason as follows: if we assume safety is necessary for knowledge,[14] might it also be *sufficient* whenever one's belief is true? And if safe, true belief is *sufficient* for knowledge, then this would make any kind of 'ability' condition whatsoever on knowledge – even a very weak ability condition – at best redundant (if safety entails that a belief derive from ability) and at worst an unnecessarily demanding extra necessary condition.

It's at this point that uni-level virtue epistemology has an important move to make, one that's been developed in different ways by Duncan Pritchard and Ernest Sosa.

The first point to note here is that at least a weak ability condition on knowledge would *not* be made redundant by a safety condition. Some beliefs that are safe are not produced from any ability *whatsoever*. The second and crucial point is that, in cases where safe true belief is not produced from ability, the belief plausibly falls short of knowledge. From these claims it follows that it's *not* the case that safe true belief is sufficient for knowledge. The argument goes as follows:

*Safety Insufficiency Argument*

1. It's possible that a true belief is both safe and not the product of ability (whatsoever).
2. True beliefs that are safe but not the product of ability (whatsoever) aren't known.
3. If safe true belief is sufficient for knowledge, then it's not the case that beliefs that are safe but not the product of ability (whatsoever) aren't known.
4. Therefore, safe true belief is not sufficient for knowledge.

The interesting premises here are (P1) and (P2). We can support both premises in one fell swoop with Pritchard's case of TEMP:

> TEMP: Temp forms his beliefs about the temperature in the room by consulting a thermometer. His beliefs, so formed, are highly reliable, in that any belief he forms on this basis will always be correct. Moreover, he has no reason for thinking that there is anything amiss with his thermometer. But the thermometer is in fact broken, and is fluctuating randomly within a given range. Unbeknownst to Temp, there is an agent hidden in the room who is in control of the thermostat whose job it is to ensure that every time Temp consults the thermometer the 'reading' on the thermometer corresponds to the temperature in the room. (Pritchard 2012, 260)

[14] This is a point we engage with critically in much more detail shortly.

The case of TEMP offers a kind of 'proof of concept' that you can have safety without aptness – and indeed (as per P1), even without the exercise of any ability whatsoever. The helper is doing *all* of the work, as it were, to make sure that Temp's beliefs are true. Moreover, the case of TEMP offers support for (P2). It seems that the disconnect between Temp's own abilities and his getting it right about the temperature in this case suffices to disqualify him as a knower, even though the helper ensures that Temp couldn't easily be wrong when looking at the broken thermometer. It follows, then, that safe, true belief is *not* sufficient for knowledge. So *even if* a safety condition is fit for the purpose of ruling out Gettier cases, knowledge can't simply be a matter of safe, true belief.

What *other* lesson can be gleaned from the case of TEMP? Does it follow that aptness is *necessary* for knowledge? It would be great for the uni-level virtue epistemologist if we *could* draw this conclusion. But we can't; the case of TEMP is a case of safety without aptness, true, but it is *also* a case where not *any* ability is present. All we're entitled to conclude from cases like TEMP is that *some kind of ability condition* on knowledge is necessary. That condition might be weaker than the kind of ability condition that aptness requires.

With this in mind, consider Pritchard's weak ability condition on knowledge:

**Weak ability condition on knowledge (WACK)**: *S* knows that *p* only if *S*'s believing *p* truly is to a significant degree attributable to their cognitive ability.

What does 'to a significant degree attributable' mean? Pritchard is careful to emphasise that the satisfaction of this weak ability condition does *not* imply (though is implied by) Greco's stronger ability condition. Remember that, for Greco, ability must be – in cases of knowledge – *the most salient* part of the total set of causal factors that explains why the subject believes truly.

We'll soon see why Pritchard thinks Greco's ability condition is too strong, and that WACK is better. But for now it should suffice to conclude first that even though safety can rule out Gettier cases just like aptness can, safe true belief isn't sufficient for knowledge. (Thus far, we've seen no reason to think apt belief is *not* sufficient.) Second, cases like TEMP motivate *at least* a weak ability condition such as WACK on propositional knowledge. This is good news for uni-level virtue epistemology; but, equally, it is good news for at least one formidable competitor: Pritchard's (2012) *anti-luck virtue epistemology.*[15]

[15] Pritchard's most recent formulation of anti-luck virtue epistemology is *presented as anti-risk virtue epistemology* (2016), which keeps most of the key details the same. Since for our purposes what goes for his anti-luck virtue epistemology will go likewise for the newer version, my presentation of the view will focus on the earlier anti-luck formulation.

## 1.6 Comparison: Anti-Luck Virtue Epistemology

Keeping score so far: uni-level virtue epistemology of the sort we've been exploring – which identifies knowledge with Greco-aptness – has the following merits; it:

(i) rules out Gettier cases;
(ii) implies a ready solution to the value problem; and
(iii) rules out knowledge in the TEMP case.

By comparison: although the view that knowledge is safe true belief can (i) rule out Gettier cases just as well as can an identification of knowledge with Greco-aptness, it has no clear explanation for (ii) and outright fails (iii).

Pritchard's (2012) *anti-luck virtue epistemology* is a more difficult competitor.

**Anti-luck virtue epistemology (ALVE)**: *S* knows that *p* iff *S*'s true belief that *p* satisfies both (i) the safety condition (SC); and (ii) WACK.

According to ALVE, SC and WACK are logically independent necessary and jointly sufficient conditions for knowledge. One knows iff one satisfies both SC (which accommodates the insight that knowledge excludes luck) and WACK (which, for Pritchard, is what is motivated theoretically by the intuition we have – for example, as illustrated by cases like TEMP, that knowledge must in some way be the product of ability).

Anti-luck virtue epistemology is tough competition because it (i) rules out Gettier cases (courtesy of the safety condition); and (iii) unlike a simple 'safe true belief' account of knowledge, rules out knowledge in TEMP-style cases, given that WACK is not satisfied in such cases. As for (ii), an explanation for the value of knowledge, the edge still goes to uni-level virtue epistemology, which identifies knowledge with Greco-aptness. *That* view, recall – in conjunction with the VOA thesis – implies that knowledge is finally valuable. However, we can't derive the thesis that knowledge is finally valuable from the conjunction of VOA and WACK. This is because there is a logical gap between WACK and the thesis that knowledge is a (species) of cognitive achievement (CAT).

One might ask whether we're being uncharitable to ALVE here. Perhaps we could interpret WACK in a way that would imply CAT (and thus, in a way that would combine with VOA to generate the result that knowledge is finally valuable)?

What's interesting here is that it is actually important for Pritchard that he does *not* close this gap! Pritchard's ALVE is *designed* in such a way that we *should not* read WACK as implying CAT. But the reason for this insistence – the