

THE CAKE THEORY OF CREDIT

forthcoming in Philosophical Topics (pre-proof)

Jaakko Hirvelä & Maria Lasonen-Aarnio

ABSTRACT: The notion of *credit* plays a central role in virtue epistemology and in the literature on moral worth. While virtue epistemologists and ethicists have devoted a significant amount of work to provide an account of creditable success, a unified theory of credit applicable to both epistemology and ethics, as well as a discussion of the general form it should take, are largely missing from the literature. Our goal is to lay out a theory of credit that seems to underlie much of the discussion in virtue epistemology, which we dub the *Cake Theory*. We argue that given the goals that virtue epistemologists and ethicists who discuss moral worth have, this theory is problematic, for it makes credit depend on the wrong facts.

1. INTRODUCTION

The central theme of Ernest Sosa's *Epistemic Explanations* is that epistemic normativity is telic. That is, epistemic normativity is the normativity of epistemic attempts as epistemic attempts, and epistemic normativity is thus insulated from non-epistemic considerations, such as whether the attempt is morally impeccable. In virtue epistemology, as in virtue ethics, the normative standing of performances is explained in terms of the agent's properties such as whether a relevant success issues from a competence or virtue seated in the agent. The most important normative standing in telic virtue theory is the fully creditable success Sosa (forthcoming, p. 24). A success is creditable to an agent just in case the agent succeeds sufficiently through competence (Sosa, forthcoming, p. 19). In Sosa's terminology, a success that is sufficiently due to competence is labeled *apt*. One of the central tenets and applications of the performance normative framework is a view of knowledge as creditable true belief.

Virtue epistemologists think of credit as *attributability*. If a success is creditable to an agent, then the success is attributable to the agent and hence, the agent is responsible for it. What is the relation of such attributability to evaluations of praise and blame? It is widely accepted that responsibility is intricately tied with being eligible for responses such as praise and blame: We can, in principle, be properly praised

and blamed only for actions and successes that we are responsible for.¹ Indeed, Sosa (forthcoming p. 163) himself says that “At the core of [telic] assessment are questions of credit and blame, and of responsibility even of the sort that involves only *attributability* and not necessarily *accountability*.”² While Sosa does not say very much about responsibility, we take his talk of responsibility (in the sense of attributability) at face value.³

Given a telic framework, it is plausible that there are performance-domain specific kinds of praise and blame, and, correspondingly, domain specific kinds of responsibility. For instance, a person can in principle be praised for a murder, *qua* perfect murder, while being morally blamed for the murder. We return to the connections between credit, attributability, praise and blame in several places below.

The notion of credit plays at least three important theoretical roles in virtue epistemology. First, and perhaps most importantly, credit is used to explain the sense in which knowledge is incompatible with belief that is merely true by luck. Virtue epistemologists see the relevant kind of luck as intricately connected with credit: the more a cognitive success is due to good luck, the less it is creditable to the agent; and the more creditable it is, the less it is due to luck (Sosa, forthcoming, pp. 19, 22). Given that in Gettier-cases the subject believes the truth in large part due to good luck (Zagzebski, 1994), virtue epistemologists draw on the notion of credit to solve the Gettier-problem.

Second, virtue epistemologists appeal to the notion of credit to explain why knowledge has added value over mere true belief. Some virtue epistemologists maintain that creditable successes constitute achievements, and that achievements are finally valuable. Since knowledge is creditable true belief, it is an epistemic achievement (Greco, 2010, pp. 97-98; Sosa, forthcoming, p. 19). Assuming achievements to have added value over mere successes, it follows that knowledge is more valuable than mere true belief. But as Sosa aptly points out, one can earn telic credit for a success without the success being in any sense valuable: “A shot might be a “perfect” murder, including its excellence *as a shot*, and *thus* creditable to its agent, while constituting an abominable crime, to the agent’s moral *discredit*” (Sosa, forthcoming, p. 20). This passage is revealing, since it opens up another way in which the notion of credit could help virtue epistemologists solve the *Meno* problem. The thought is that by deserving

¹ See, for instance, Smith (2015, p. 106).

² Though he does not say so, we suspect Sosa has in mind something like the distinction between two senses of responsibility (attributability and accountability) made by Watson (1996).

³ J. Adam Carter notes in correspondence that telic assessments are applicable even to lower-level animals which intuitively fall outside appropriate responsibility attributions, and that this complicates the relationship between responsibility and telic credit. First, we would not rule out the possibility that a lower-level animal could be responsible (in the sense of attributability) for events. Second, most of our criticisms target the idea that telic credit is necessary, though perhaps not sufficient, for responsibility, an idea which we assume Sosa to endorse.

credit for attaining an aim one in some way taps into the value of that aim. If the aim itself is commendable, then the credit for the success that one deserves has a positive valence - for instance, a morally positive value. If the aim is deplorable, the credit has a negative value. While truth is sometimes worse than useless from a practical point of view, it is not uncommon to think that it is nevertheless finally valuable. Thus, the added value that knowledge has over mere true belief would be achieving something finally valuable through one's cognitive agency. Note, however, that this way of solving the value problem, unlike one appealing to the final value of achievements, is premised on the value of truth.⁴

Third, the notion of credit plays a unique role in Sosa's virtue epistemology. Sosa holds that knowledge comes in three varieties: *animal knowledge*, *reflective knowledge (full well)*, and *secure knowledge (full well)*. The hierarchy of these different grades of knowledge is ordered in terms of the degree to which a success (truth or aptness) is creditable to the agent. In the case of animal knowledge one believes truly through a first-order competence and hence, is creditable for believing truly. But animal knowledge is still susceptible to a kind of credit-reducing luck, since being creditable for believing truly - at least to the extent required for animal knowledge - doesn't entail being creditable for believing aptly. In the case of reflective knowledge full well the aptness of the true belief is due to a meta-competence of the agent, such that she is sensitive to when her beliefs would be apt, and not just true. As a result, the agent is credited not only for believing truly, but also for believing aptly. On the ultimate level of secure knowledge one attains reflective knowledge full well, and the attainment of this knowledge derives from competences that are retained securely, and not just accidentally. One has secure knowledge only if one is safe from losing one's competence to judge correctly while retaining a disposition to judge. If one's knowledge is unsecure one could easily have formed a belief without possessing the relevant epistemic competence (Sosa, forthcoming, pp. 161, 175-176). We return to this hierarchy in §5: as we will see, Sosa's thought seems to be that as one ascends the knowledge hierarchy, it's not just that one deserves credit for different successes, but that one deserves more credit for one's *true belief*.

Though the two literatures have been for the most part isolated, it is worth noting that the notion of credit also plays an important role in the literature on moral worth. The following thought is widely accepted in that literature: a morally right action is morally worthy just in case the agent deserves sufficient credit for doing the right thing. This points to a parallel between virtue epistemological views of knowledge and views of moral worth. Just like knowledge, moral worth requires success (doing the right thing) that is not merely lucky or accidental. If the agent deserves sufficient credit for doing the right thing, then it is not a mere accident that they did the right thing. Hence, like virtue epistemologists,

⁴ Baehr (2012) argues that some true beliefs lack any positive value and hence this way of explaining the added value of knowledge doesn't pan out.

those writing on moral worth contrast creditable success (morally right action) with success that is lucky or accidental.⁵

But how should we understand credit itself, and is there a single notion that could serve us both in epistemology and in ethics? In §2 we lay out one way of thinking about the structure of credit, which we call the Cake Theory of credit, that seems to underlie much of current virtue epistemic theorizing. In §3-§5 we examine different ways in which virtue epistemologists could endorse the central tenets of the Cake Theory - in particular, different views of what grounds degrees of credit. In §6 we argue that the Cake Theory is problematic and consider possible replies on behalf of virtue theorists. In §7 we conclude by laying out some desiderata that theories of credit ought to satisfy.

2. THE CAKE THEORY OF CREDIT

In this section we formulate the bare bones of a theory of credit that seems to underlie much of virtue epistemic theorizing.

⁵ The idea that the kind of luck at issue is the same in the two cases is at least worth exploring, due to the theoretical unification that could be achieved. Sliwa (2016), for instance, understands accidentality in the moral domain as the kind of accidentality that is incompatible with knowing. Note further, that the usefulness of a unified account of credit and attributable success across ethics and epistemology is not premised on a view of knowledge as creditable true belief. One might, for instance, hold that though the relevant success in ethics is doing the right thing, in epistemology it is knowledge, not true belief. On such a view one might know, even if the knowledge is not attributable or creditable to the knower. (Indeed, both of the authors have defended views along these lines; see Hirvelä (2018, 2019b) and Lasonen-Aarnio (2010, forthcoming).

Virtue epistemologists understand competences as dispositions to succeed while in certain kind of circumstances.^{6, 7} An epistemic competence, then, is a disposition to attain or retain true beliefs.⁸ Virtue epistemologists have talked about success being “owed to”, “due to”, “in virtue of”, “because of” one’s competence, or of one’s competence being “responsible” for the success.⁹ For now we leave it open just what relation between one’s competence and the relevant success such talk picks out, and will use “responsible for” to pick out the relation in question. Whatever the relation is, other factors (other agents, environmental factors such as random gusts of wind, etc.) can also be responsible for the success to various degrees.

According to the Cake Theory, whether a success is creditable to an agent, full-stop, depends on the extent to which the agent’s competence, as opposed to other factors, is responsible for the success: the larger the role played by these other factors, the more the success is due to luck. Think of the relevant kind of responsibility here as a cake: in any given case, there is only a limited quantity to be had. Different factors responsible for the success are assigned pieces of the cake, the sizes of which correspond to the degree to which they are responsible for the success. For an agent to be creditable for the success, full-stop, her competences must get a big enough slice of the responsibility-cake. The core of the theory, then, is the thought that the degree to which a successful performance is creditable to an

⁶ The idea that competences are dispositions to succeed is widely shared among virtue epistemologists (Greco, 2010; Littlejohn, 2014; Pritchard, 2012). Miracchi (2014, pp. 44-46) is a notable exception since she considers competences to be a *sui generis* property that only subjects (in virtue of having cognitive systems) can have, which allow the subject to reliably attain success when exercising the competence in question. Vetter (2016) argues that the dispositional account of competences invoked by virtue reliabilists does not give a general account of competences, since abilities and dispositions have a different modal structure.

⁷ Sosa understands dispositions to be properties characterisable by sets of trigger-manifestation conditionals, called SSS-conditionals. An SSS-conditional is a conditional of the form: if X were to host a seat Se of a disposition D, while in an appropriate situation Si and shape Sh, and D were triggered, then X would ϕ . X property P is a seat of a disposition D only if X having P would make the relevant SSS-conditionals true of X. To have a disposition D, X must have some seat Se of D. Shapes are intrinsic properties. Situations are extrinsic properties. The appropriateness of shapes/situations is determined by community interests. (Sosa, 2015, p. 104). Hirvelä and Paterson (2021) argue that neither knowledge nor intentional action requires that one possesses competences as Sosa conceives them.

⁸ Kelp (2017) and Miracchi (2015) hold that the relevant epistemic success is *knowledge* rather than true belief, and hence understand competences as dispositions or propensities to achieve knowledge. Hirvelä (2019b) argues that one can know without having the competence to do so.

⁹ This placeholder use of “responsible for” should be distinguished from a more substantive sense of “responsibility” - indeed, one at play when Sosa cashes out credit in terms of a kind of responsibility - that we will also discuss below.

agent is a function of how responsible her competence is for the success, compared with other factors: degrees of credit are a function of the proportion of the cake assigned to her competence.

The Cake Theory of Credit

- (1) Credit comes in degrees: one can deserve more or less credit for a success.
- (2) The degree to which a success is creditable to an agent A is a function of the degree to which A's competence, as opposed to other factors, is responsible for the success.
- (3) One needs a sufficiently high degree of credit for a success in order to be creditable (full-stop) for the success.

A view along these lines appears to be widely endorsed by virtue epistemologists. Consider, for instance, the following passages:

The success of an attempt is creditable only if sufficiently owed to the agent's pertinent triple-S competence. (Sosa, forthcoming, p. 119)

How much a success is by *credit-reducing* luck depends on how little it is due to competence. *Excess of such luck* aligns with *deficiency of corresponding competence*, which reduces or blocks relevant "credit" to the agent for the success of their attempt. (Sosa, forthcoming, p. 22)

[In cases of knowledge] the person derives epistemic credit ... that she would not be due had she only accidentally happened upon a true belief ... The difference ... here is the variation in the degree to which a person's abilities, powers, and skills are causally responsible for the outcome, believing truly that *p*. (Riggs, 2002, pp. 93-94, as cited by Greco 2010, p.72)

When we attribute knowledge to someone we imply that it is to his credit that he got things right. It is not because the person is lucky that he believes the truth — it is because of his own cognitive abilities. He figured it out, or remembered it correctly, or perceived that it was so. (Greco, 2003, p. 123)

But why exactly are virtue epistemologists committed to the Cake Theory?

Virtue epistemologists think that knowledge is essentially creditable true belief, and agree that beliefs true by luck cannot constitute knowledge.¹⁰ As we have seen, virtue epistemologists, along with many

¹⁰ Many hold that a belief is not true by luck only if it is safe from error in that it could not easily have been false (Hirvelä, 2019a; Lasonen-Aarnio, 2010; Pritchard, 2005). Several virtue epistemologists have argued that

writing in the moral worth literature, hold that the more a relevant success is due to luck, the less creditable it is to the agent, and the more creditable it is to the agent, the less lucky it is. Since a success can be more or less lucky, a success can be more or less creditable to an agent. Credit must therefore come in degrees. Therefore virtue epistemologists ought to endorse (1).

As we have seen, virtue epistemologists accept the schematic idea that when a success is creditable to an agent, the agent's competences are responsible for the success. Moreover, recall the opposition between credit and luck: the more credit, the less luck; the more luck, the less credit. Already inherent in the opposition between credit and other, luck-constituting factors, is the idea that one's competences compete with these other factors. If we add the assumption that the degree to which a success is lucky is a function of the degree to which factors other than one's competence are responsible for the success, (2) follows.

It should already be clear why virtue epistemologists accept (3). If a success is in large part due to luck, one can earn some credit for the success, but not enough for it to be creditable, full-stop. For instance, virtue epistemologists think that knowledge is creditable true belief and excess of good epistemic luck is incompatible with knowing: earning only a minimal degree of credit for one's true belief entails that it is too lucky that one believes truly, Gettier-cases being a case in point (Lackey, 2007). Similarly for other successes.

We will now examine different ways in which virtue epistemologists could endorse the central tenets of the Cake Theory - in particular, different views of what it takes for an agent's competences to be responsible for a success.

3. DEGREES OF CREDIT AND CAUSAL RELEVANCE

Greco (2003, 2010) and Riggs (2002, p. 91) hold that the degree of credit an agent deserves for a success is a matter of the causal relevance of the competences of the agent: the degree to which a success is causally determined by the competences of the agent (Riggs, 2002), or how salient a factor the competences are in a causal explanation of the success (Greco 2010). On this view, the degree of creditability of a success to an agent is a direct function of the causal relevance or salience of the competences exercised by the agent to attain the success.

satisfying a virtue theoretic condition entails satisfying a safety condition (Carter, 2016; Gaultier, 2014; Greco, 2016; Littlejohn, 2014). Sosa's relationship to safety has been complicated, but nowadays he holds that knowledge entails a kind of safety condition (Sosa, 2015) (see (Greco, 2016) for discussion). However, one of us has argued that satisfying the virtue theoretic condition doesn't entail satisfying the safety condition (Hirvelä, 2020; Hirvelä & Paterson, 2021). See also Pritchard (2010).

Causal relevance is one way of cashing out the kind of responsibility in terms of which the Cake Theory was formulated: one's competences and other factors compete for causal relevance, and one's competence must get assigned a big enough slice of the relevance cake for the success to be creditable. Note that several substantial assumptions are being made here. For one thing, it is assumed that different causally relevant factors share a common currency in which their relative contributions can be evaluated. But as noted by numerous philosophers of science, such an assumption of a common currency is problematic (see e.g. Sober (1988)).

Perhaps Greco's view, appealing not to an objective measure of causal relevance, but to the explanatory salience of one's competence, bypasses these worries by avoiding the need for objective comparisons of causal powers or relevance (Greco 2003). But causal explanations are notoriously context sensitive. The police might deem a car crash to have been caused by excessive speed while city planners might hold it to have resulted from difficult traffic patterns (Greco 2010, p. 106). Moreover, both the police and the city planners might be correct since they operate in different contexts. Greco (2010) is explicit that his way of understanding credit leads to contextualism about knowledge. But note that as a general way of cashing out the Cake Theory of credit, such contextualism goes far beyond knowledge. If we were to apply Greco's view to morally worthy action, for instance, we would have to conclude that whether or not an action is morally worthy depends on whether the agent's competence strikes an evaluator as a particularly salient part of a causal explanation of her success of doing the morally right thing. Praiseworthiness for doing the right thing would itself become highly context-dependent - a conclusion we don't find altogether palatable.

Setting aside contextualism, Greco's view is susceptible to other, perhaps more serious, worries. It is a version of the Cake Theory: different causal factors compete for explanatory salience, and one's competences must be a particularly salient factor. It is natural to understand the view as one on which one's competences must be the most salient factor, being at least assigned a bigger slice of the salience-cake than other factors, if not bigger than all the others combined. Below we discuss problems created by joint action, but at this point it is helpful to bring up an important line of objection to the idea of knowledge as creditable true belief, due to Jennifer Lackey (2007). Consider Lackey's case *Chicago Visitor*:

Having just arrived at the train station in Chicago, Morris wishes to obtain directions to the Sears Tower. He looks around, approaches the first adult passerby that he sees, and asks how to get to his desired destination. The passer-by, who happens to be a Chicago resident who knows the city extraordinarily well, provides Morris with impeccable directions to the Sears Tower by telling him that it is located two blocks east of the train station. Morris unhesitatingly forms the corresponding true belief. (Lackey 2007: 352)

Morris acquires knowledge through testimony. But as Lackey points out, what explains why Morris has a true belief regarding the whereabouts of the Sears Tower

has nearly nothing of epistemic interest to do with him and nearly everything of epistemic interest to do with the passerby. In particular, it is the passerby's experience with and knowledge of the city of Chicago that explains why Morris ended up with a true belief rather than a false belief (Lackey, 2007, p. 194).

The worry is that given a somewhat normal context, the explanatory contribution of Morris's competence to his success of truly believing is not particularly big or salient and hence, on views like Greco's, Morris cannot be credited for his true belief. Yet, we seem to be dealing with a rather paradigm case of knowledge through testimony. Lackey considers a revised theory on which all that is required for credit is that one's cognitive faculties make some explanatory contribution, but points out that the resulting view is too weak, for the condition in question holds in Gettier-cases. Hence, the virtue epistemologist faces a dilemma: either they cannot adequately deal with Gettier-cases, or they cannot accommodate testimonial knowledge. Something has to give.

Lackey presents the Morris case as a problem for a view on which knowledge is creditable true belief, but we think that analogous cases make trouble, in effect, for the Cake Theory of Credit.¹¹ Consider the following case:

Tokyo Metro

Aino is visiting Tokyo for the first time, and has promised to meet a local friend at the Shibuya station at noon. In order to get to the station, Aino must switch metros numerous times. Knowing about the complexity of the metro network, Aino leaves ample time to get to Shibuya. Once at the station closest to her, she finds herself unable to figure out how to get to Shibuya on the basis of the metro maps she sees, so she consults a friendly local for advice on which metro to get on first. Over the course of the next 90 minutes, Aino consults over ten friendly locals, and finally finds herself at Shibuya, in ample time for the meeting.

We think that Aino is creditable both for the success of making it to Shibuya station in time, and for keeping her promise. For instance, assuming that keeping her promise is the right thing to do, she not

¹¹ We should flag that Lackey says not only that it is facts having to do with the testifier that explain Morris's true belief, but that the truth of the belief is not creditable to him. For our dialectical purposes here we are treating the case as a problem not for the view that knowledge is creditable true belief, but as a problem for a particular way of thinking about credit. Indeed, though neither of us thinks knowledge is creditable true belief, we don't think *Chicago Visitor* is a counterexample to such a view.

only does the right thing, but her action has moral worth. She can be praised for keeping her promise. Indeed, we don't think there is any relevant difference between how creditable Aino is for keeping her promise and how creditable she would have been had she walked to the station using a map, doing all the navigating on her own.

Might someone respond, on behalf of Greco, that while Aino is not creditable for making it to the station in time, she is nevertheless creditable for keeping her promise, even if the latter requires making it to the station in time? Consider explaining why Aino succeeded in keeping her promise, which was meeting her friend at the station at noon. It is not difficult to imagine a context in which the contributions of the numerous helpful locals are highly salient; indeed, it is largely due to their helpful advice that Aino managed to keep her promise. But even within such a context, we don't see these facts about explanatory salience as detracting from moral worth.

We return to cases like *Tokyo Metro* below. But a hypothesis already emerging from the discussion so far is that factors other than one's competence being responsible for a success to a high degree need not detract from the credit one deserves for the success.

4. CREDIT AND MANIFESTATIONS OF COMPETENCE

We now turn to consider an alternative way of understanding the 'responsible for' -relation. According to the manifestation account, an agent's competence is sufficiently responsible for a success if and only if the agent's competence is *manifested* in the success (Sosa, 2009, p. 12; forthcoming, pp. 60, 82, 150; Turri, 2011, p. 7). Understanding creditability in terms of manifestation of competence rather than in terms of causal relevance might have significant benefits. First, while Morris' competences are not the most salient causal explanation as to why he acquired a true, rather than a false belief, his competences are manifested in his acquiring a true belief, and hence we can maintain that his cognitive success is creditable for him. Second, virtue epistemologists have argued that the true beliefs of Gettiered subjects do not manifest their competences, so proponents of the manifestation account may be able to escape Lackey's dilemma (Turri 2011, p. 11). Third, understanding credit in terms of manifestation is elegant, since competences are dispositions and dispositions can be manifested. It is at least initially plausible to think that a disposition is responsible for an event if that event is a manifestation of the disposition. Finally, the manifestation account is not committed to contextualism about credit.

But one immediate problem for the manifestation account is that it is no longer clear how it can deliver the idea that credit comes in degrees. After all, manifesting a disposition is an on/off matter. Either, one might think, a shattering is a manifestation of the fragility of a vase or it isn't. In response, one could try to argue that dispositions can be manifested to different degrees: for instance, that when stuffed with

dynamites, the shattering of the vase manifests its fragility to a lesser degree than when it is dropped on a hard surface in somewhat normal conditions. Bradford (2015, p. 110) seems to hold this view since she writes that “[o]ne might say that the [success] manifests the competence to the degree that the competence is responsible for the [success] (although this is by no means a strict definition).” But note that here Bradford helps herself to the notion of competence being responsible for a success in explaining degrees of manifesting competence, rather than vice versa: what the virtue epistemologist would need is an independent account of degrees of manifesting dispositions. Indeed, it is hard to see how this view differs from the causal account, which basically holds that one deserves credit for a success if one’s competences are causally responsible for the success.

Hence, it is worth noting at the outset that *prima facie* at least, the manifestation account is at odds with the Cake Theory, which assumes credit to come in degrees. While the Cake Theory, and especially the idea that credit comes in degrees, might be hard to reconcile with the manifestation account, we nevertheless take Sosa to be committed to both, as evinced by the quotations given above. We turn to explore how Sosa can consistently hold on to both ideas.

In the past Sosa has advocated the idea that the degree of credit one earns for a success depends on how reliable the manifested competence is: “Aptness comes in degrees. One shot is more apt than another, for example, if it manifests a more reliable competence.” (2009, p. 13) But what does the reliability of a competence amount to? Note at the outset that we cannot understand the reliability of a competence as a function of the range of circumstances in which its manifestations are successful - for instance, roughly as how often an archer’s competence, when manifested, issues in a successful shot. That is because Sosa characterizes competences as dispositions to succeed. But a disposition to succeed cannot have unsuccessful manifestations, any more than a disposition to break can have as its manifestations non-breakings. As far as we can see, there are two remaining ways in which the reliability of a competence could be understood.

First, a disposition could be said to be reliable to the degree to which it is likely to manifest when triggered. That is, setting aside the phenomenon of masking, the higher the chance of the manifestation of the disposition is when the stimulus condition is present, the more reliable the disposition is. It would follow that if a success manifests a so-called surefire disposition, then the degree of credit one deserves for the success is maximal.¹² But the resulting view is implausible, for it seems to entail that the degree of credit that can be earned for extremely difficult performances is lesser than the degree of credit that can be earned for easy performances. After all, difficult performances plausibly often involve deploying competences that are less reliable in the sense under discussion. Suppose I have both a competence to win in chess against opponents whose rating is 800 and against opponents whose rating is 1600. The

¹² For a distinction between surefire and probabilistic dispositions, see Prior, Pargetter, and Jackson (1982).

latter competence is less reliable in the relevant sense, for it is less likely to manifest when triggered. Nevertheless, I need not deserve less credit for winning an opponent whose rating is 1600 than I deserve for winning an opponent whose rating is 800 (indeed, the two games may involve identical moves). More generally, a probabilistic disposition can be fully responsible for its manifestation, since no external factor need contribute to its manifesting. Think for example of an isotope of uranium 238, which has a half-life of roughly 4,5 billion years. Within any relatively short interval of time the probability that it manifests its disposition to decay is extremely low. Yet, when it does so its disposition to decay is fully responsible for its decay, since decaying is a spontaneous process, intrinsic to the isotope.

Second, the reliability of a disposition might be said to be a function of the range of possible circumstances in which it issues a manifestation when triggered. The wider the range, the more reliable the disposition is. But we think the resulting account of degrees of credit is implausible. Suppose that both you and I are competent archers, both in daylight and during the night. We are now shooting in broad daylight. However, you are more reliable than me at night, due to having better night vision: you would hit the target over a wider range of circumstances in which it is dark. If both of us manifest archery competence in hitting the target at midday, it doesn't seem that your success is more creditable to you in virtue of the fact that you are disposed to succeed in a wider range of night-time circumstances than I am. While you plausibly are the better archer, that doesn't matter, since we are interested in assessing the token attempt, not whom of us is better at archery.

We doubt that the degree of credit one earns for a success could depend entirely on general features of one's competences, such as how reliable they are - in either of the senses discussed above. To wrap up, it isn't clear how the manifestation account could make sense of the idea that credit comes in degrees, but it is clear that Sosa is committed to both the manifestation account and to the idea that credit comes in degrees. In the next section we consider Sosa's latest account of degrees of credit.

5. DEGREES OF CREDIT AND SOSA'S HIERARCHY OF ATTEMPTS

Recall that Sosa holds that there are three levels of knowledge: the animal, reflective and secure knowledge. On the animal level the subject's true belief manifests her competence to believe the truth. On the reflective level the subject manifests a meta-competence to judge well, and gains credit for the aptness (not just the truth) of her first-order belief. On the level of secure knowledge the subject attains reflective knowledge through competences retained safely. But it's not just that one is creditable for different things – for instance, in the case of animal knowledge, for true belief, and in the case of reflective knowledge, for apt belief. Sosa's thought seems to be that as one ascends up the hierarchy, one deserves more credit for one's *true belief*:

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. (Sosa *forthcoming*, p. 176)

We take the view to be that the higher up the hierarchy I ascend, the more credit I deserve for the truth of my belief, and the less lucky (in the telically pertinent sense) it is that my belief is true. Full, luck-excluding credit requires that one's knowledge be secure.

Given Sosa's performance normative framework, this tripartite hierarchy of knowledge is an instance of a general hierarchy of attempts: an attempt can be apt, reflectively apt, or secure. And so the general thought is that the degree of credit one deserves for a success depends on how high one ascends this hierarchy: to attain full credit for a success, one's success must be secure. If we take Sosa's idea that credit is tantamount to a kind of responsibility (in the sense of attributability) at face value, it follows that one is *more* responsible for a success – it is more attributable to one, more one's own doing – the higher up one's attempt places in the hierarchy.¹³

Let's look a bit more closely at secure attempts, which is the newest element in Sosa's framework. According to Sosa having a complete competence (a so-called SSS-competence) requires having a seat of the competence while being in the right kind of shape and situation (see footnote 7 for more details). On Sosa's view, the attainment of secure performance requires that one's SSS (Situation, Shape, Skill) -competence is safe in the following way: one could not easily have failed to retain one's SSS-competence, while still making the same (or, perhaps, a relevantly similar) attempt. Hence, security does not require that the retaining of one's SSS-competence be safe. For instance, if one's life is in danger, then the retaining of one's SSS-competence is, of course, in danger. Yet, a pilot in grave danger of being shot down can come to have secure knowledge: they are safe from losing their SSS-competence while still retaining a disposition to make judgments regarding a relevant question even in the absence of the complete competence (Sosa *forthcoming*, p. 161). The contrast case considered by Sosa is that of Simone, a fighter pilot who, unbeknownst to her, is routinely tested under simulation. Simone, Sosa

¹³ Even if credit is a kind of responsibility, we are not sure how Sosa's framework could be extended into a general theory of responsibility. I can be responsible for doing something badly, for making a mistake, or for an omission – for instance, for forgetting to answer an important email. Incompetent actions and omissions can, it seems, be attributable to me, but they are not manifestations of competence. Many thanks to Ninni Suni for discussion.

assumes, could easily have been under such a simulation, in which case the situational component of her SSS-competence would have been missing. When she judges, for instance, that she is flying over a lake, she lacks secure knowledge, for she could easily have made this judgment even while lacking her complete SSS-competence. While Simone can have animal and reflective knowledge, she lacks secure knowledge.

Should we agree, then, that a success becomes more creditable to a subject in the sense of being more attributable to her in virtue of the fact that it is not only apt, but reflectively apt and secure? Our worry is that the resulting view makes degrees of credit – and hence, degrees of responsibility in the sense of attributability – dependent on the wrong kinds of facts. We think an agent who retains her competences in an insecure way can be fully responsible for her actions. Consider the following example:

Poisoned Apple

Malla has decided to kill Saana by giving her a poisoned apple. But she must first mix the poison. To manufacture a potent enough poison, she must mix 10 ingredients in the right quantities. She sets out to do so, putting in just the right number of drops of each ingredient using a pipette. However, the lights in the room are highly unreliable, occasionally gradually dimming in a slow, imperceptible manner. Had the lights been dimmer, Malla would not have been able to count the right number of drops of each ingredient – as a result, the quantities would not have been precise enough, and the concoction not potent enough to kill Saana. But as it happens, the lights work perfectly, and Saana is poisoned by Malla's apple.

In this case, Malla does not safely retain the situational component of her complete SSS-competence to poison Saana: there are modally nearby cases in which she attempts to poison Saana even though she lacks the complete competence to do so. Yet, this doesn't make Malla any less responsible for poisoning Saana – poisoning Saana is no less Malla's own doing. How reliable the lights are in the room where Malla mixes her poisons just seems irrelevant when thinking about issues of attributability and responsibility here. And as we saw at the outset, responsibility, in the sense of attributability, is plausibly connected to attributions of praise and blame: if an agent is not responsible for an action, then she cannot properly be blamed or praised for the action, since it wasn't really their doing. But Malla is no less blameworthy because she could easily have botched up her concoction due to malfunctioning lighting.

Note that these points are not confined to cases involving responsibility for actions that are morally wrong. Consider the following case:

Depth Perception

Halti sees a child drowning in a river. Luckily there is a lifebuoy by the bank that she can throw to the child, though to throw it far enough she must precariously balance herself in a dangerous place above some sharp rocks. Halti does so, successfully throwing the buoy to the child, who then grabs it and is saved. However, Halti's eyesight is very sensitive to moisture: when the air is as moist as it is here, right by the roaring river, her depth perception is severely affected in about 50% of cases. Halti cannot tell when it is and isn't affected. Had her eyesight been affected, she would have misjudged the location of the child, throwing the buoy in the wrong place.

Consider Halti's competence to save the child, which essentially involves a competence to act quickly enough, throwing the buoy to the child. Her SSS-competence is not safe, for had her eyesight been affected, she would have lost the shape component of her competence to throw the buoy. Nevertheless, she would still have made an attempt to save the child, throwing the buoy at the location she would then have falsely believed to be that of the child. Does the fact that her eyesight is precarious in this way make her any less creditable or responsible for saving the child? Does it make her less praiseworthy; does it make her action less morally worthy? We think not. Halti is every bit as creditable for saving the child as someone with no such moisture-sensitive defect in their eyesight; indeed, the defect in question just simply seems irrelevant for issues of credit, responsibility, praise, and moral worth.¹⁴

The reader might here demur at our talk of moral responsibility and worth, given Sosa's focus on telic normativity. Even though Sosa is clear that the kind of credit he is interested in is a matter of responsibility in the sense of attributability, the responsibility at issue, as such, is not moral: one can deserve telic credit for morally good, neutral, or bad actions. Could Sosa reply that he is only interested in a kind of responsibility that is *not* moral – and, perhaps, that the cases we've been pressing seem convincing because they elicit moral evaluations? Whereas moral responsibility is intricately tied with our practices of morally praising and blaming agents, there are also practices of praising and blaming that are not moral. Indeed, a coach's praise on one of her players need not in any way be moral. Similarly, one might think that as one ascends up Sosa's knowledge hierarchy, different grades (and possibly kinds) of epistemic praise are appropriate.

¹⁴ Many theories in epistemology relativise knowledge or justification - and epistemic assessment more generally - to a basis or ways of believing: we are interested in evaluating how well the subject does given that she believes on a given basis, whether thought of in terms of the evidence or reasons the belief is based on, or the method deployed. Whether the subject would have believed something false, for instance, had her belief been differently based, seems simply irrelevant. For Sosa the relevant ways of believing are a matter of which competences are manifested in believing. By requiring security the assessment of the performance is no longer constrained to considering cases where the 'basis' is held fixed.

Credit in Sosa's sense is not, of course, tantamount to moral responsibility. Nevertheless, we think it would be theoretically disappointing to wholly detach moral responsibility – which is necessary for morally crediting, praising, and blaming – from other kinds of responsibility. After all, we need to say what unites different kinds of responsibility – what makes them all kinds of *responsibility* (at least in the sense of attributability). Perhaps even more pertinently, it is not clear if Sosa's axiological framework *can* be thus detached from the realm of moral evaluation, given its general nature (often also taken to be a main selling point). After all, in evaluating Halti our interest may be in how creditable Halti is for the success of *doing the morally right thing*. And on Sosa's virtue-theoretic picture, this will be a matter of Halti's competences bearing an appropriate relation to her moral success (e.g. of the success manifesting her competence). But isn't being creditable for doing the right thing exactly what morally worthy action is?

Another possible reply to our argument might be that we are rarely concerned with the degree of credit afforded by security: aptness and reflective aptness are almost always enough.¹⁵ Yet, for security not to be an idle dangler with little theoretical work to do, we should be able to imagine situations in which precarious eyesight makes a difference for credit and responsibility. But we have a hard time imagining such cases. Even more importantly, we think a theory of moral worth, for instance, on which such factors could make a difference would be wrong: a practice that praises agents with stellar eyesight more for saving children would simply be misguided, being sensitive to factors irrelevant for such moral evaluations.

Sosa suggests a view of degrees of credit appealing to a hierarchy of attempts. The higher one ascends up the hierarchy – from *apt* to *reflectively apt* and *secure* – the more credit one deserves for a success. We have argued, however, that the resulting view makes degrees of credit depend on the wrong kinds of facts. For now we set these considerations aside and accept for the sake of argument that degrees of credit can be understood within the manifestation account. In the next section we will argue that the Cake Theory has problematic consequences.

6. PROBLEMS FOR THE CAKE THEORY

A problem with the Cake Theory that we have already gestured toward is the way in which it divides factors responsible for a success into two groups: an agent's competence, and everything else. All factors belonging to the second group are treated on a par, as constituting a kind of luck that detracts from the attributability of a success. As we see things, this is the reason why cases of testimonial knowledge are troublesome for virtue epistemological theories of knowledge that assume the Cake Theory. In such cases credit for true belief has to be shared, and often the testifier is more responsible

¹⁵ See Sosa's remarks (forthcoming, p. 171).

for the truth of the belief than the receiver of the testimony. But sharing epistemic credit with someone with whom you cooperate in order to know, as in testimonial cases, is not incompatible with knowing: a good testifier does not render the truth of one's belief lucky. The Cake Theory is insensitive to this difference.

One initially promising way to respond to this problem is to claim that in cases where one gains knowledge through cooperation credit is not shared in the first place, since the success is not attributable to any single agent, but to a cooperating group of agents. Indeed, Greco (2018, pp. 280-281) has argued that testimonial knowledge should be understood as a *joint achievement* that is attributable to the *joint agency* partially constituted by the speaker and the hearer. Greco holds that cooperative knowledge generation, say within a research team, also involves joint agency. The guiding idea here seems to be that if a success is attributable to joint agency, in which one competently participates in, then the contributions of different agents don't compete for credit, and hence cases of testimonial knowledge are not problematic. Sometimes a success is creditable to an individual agent and sometimes to a joint agency.

Greco rightly recognizes that such a view requires rejecting the idea that knowledge is true belief attributable to the cognitive abilities of the agent. There are, Greco thinks, two ways of knowing: "First, one may come to know by means of one's individual competent agency. Second, one may come to know by means of one's competent participation in competent joint agency. In the second case, it is important to note, one's having a true belief is attributable to the *competent joint agency*, as opposed to one's competent participation in that joint agency" (Greco 2018, p. 281).

However, the actions of two agents who work in tandem are not trivially manifestations of joint agency. Drawing on the work of Searle, Bratman and Gilbert, Greco (2018, p. 280) holds that joint agency involves at least four features:

- (i) a *we intention*, that is, an intention to participate in some activity and understand it as something they do together,
- (ii) the sharing of some set of sub-plans,
- (iii) interaction, in that what one member does affects what the others do, and,
- (iv) interdependence, in that what is done together cannot be done alone, at least in the same way.

We think that an agent can gain a sufficient degree of credit for a success even in cases in which another agent is responsible for the success to a high degree, and the above conditions for joint agency are not satisfied. Consider for instance the following case:

Paternalism

There's a lot of misleading information about the origin of Covid-19 available. You know that your child is keen to find out how Covid-19 emerged, and that she would easily be misled by false conspiracy theories. Without telling your child, you block all the sites where she could encounter misleading information about Covid-19. If you hadn't blocked the bad sites, your child would almost certainly have ended up believing falsehoods. But thanks to your pre-emptive actions, she encounters only genuine information, and consequently comes to know that Covid-19 has a natural animal origin.

You are for a large part responsible for your child's true belief, and it is clear that there is no we-intention in this case. Your child might even explicitly believe that you didn't meddle in her investigation. Yet your paternalistic actions don't rob your child of knowledge. While appeal to joint agents might help the virtue epistemologist deal with standard cases of testimony, it doesn't seem to get to the gist of the problem.

Virtue theorists might reply that the notion of joint agency is too strict, and that competences can be possessed by collectives that don't constitute joint agents. Sosa (2007, pp. 93-94) gestures towards this kind of view. Testimonial knowledge might be gained from manuscripts written long ago, and in such cases the collective responsible for the knowledge is not organized and its members certainly don't have the required kind of we-intention, since some parties might be dead. In some such cases, though individuals only deserve partial credit for a success, the credit might be enough at least for animal knowledge (Sosa 2007, p.97).

But if nothing like genuine joint agency is required, then our worry is that agents who have no intention of helping someone in F-ing can be partially responsible for the fact that someone successfully F's. Suppose that instead of trying to keep her promise in *Tokyo Metro*, Aino's aim is to blow up Shibuya station, which she succeeds in doing. If no shared intention is required in order to share credit for a success, then the friendly locals who provided Aino with directions to the station appear to get at least partial telic credit for the horrendous deed: Aino's success appears to be at least partially attributable to them. Moreover, recall that Sosa thinks of such attributability in terms of responsibility. But we think it is implausible that the locals are responsible for the blowing up of the station. Moreover - recalling the connection between responsibility and blame - the friendly locals surely cannot be blamed for the attack on the station, and neither do they need an excuse for having helped Aino.

Instead of invoking the idea of joint agents or looser collectives, virtue reliabilists might suggest - drawing on the idea of joint manifestations of dispositions - that credit is fully attributable to all the contributing parties if the success is the *joint manifestation* of the competences of the contributing

parties.¹⁶ The idea would be that two dispositions can jointly manifest themselves in a single success, and in such cases both can be fully (or at least sufficiently) responsible for the success. It does not follow, however, that these two dispositions must constitute a competence (had by some sort of collective). (Figuring out whether the resulting view is compatible with the Cake Theory is something we leave as an exercise to the reader.)

The joint manifestation account may be able to accommodate *Paternalism*, assuming that a case could be made for thinking both your and your child's competences are jointly manifested in the true belief of your child. It delivers the right verdict in the version of *Tokyo Metro* where Aino blows up the station, since plausibly the passerby's competences are not manifested in Aino's blowing up of the station, though they may be manifested in her arriving at the station in time. It may also be able to give the correct verdict in the following kind of case:

Methanol and Ethanol

While hiking in the woods Olos encounters a man who suffers from severe methanol poisoning. Olos knows that methanol poisoning is treated by ingesting ethanol. Olos has a pocket mat with him and gives it to the man. But he knows that the amount of ethanol it contains is not sufficient, so after giving the man the pocket mat he runs for help. A few minutes later Ounas, who knows nothing of Olos, encounters the poor chap, recognizes that he suffers from methanol poisoning, and gives him the small dose of ethanol she carries in her pocket mat. But Ounas knows that the amount of alcohol she had isn't sufficient to cure a severe case of methanol poisoning, and runs for help. Neither Olos nor Ounas can find any help, but their combined effort cures the man's methanol poisoning, and he doesn't lose his sight.

In *Methanol and Ethanol* Olos and Ounas deserve equal credit for curing the man, and intuitively their actions are equally morally worthy. Moreover, perhaps we can view the success (curing the man) as a joint manifestation of their relevant competences. So the joint manifestation account might be able to deliver the right verdict here. But alter the case slightly and it delivers the wrong verdict. Hold everything about Ounas fixed, but assume that Olos is incompetent and knows nothing about how to cure methanol poisoning. Suppose, for instance, that he falsely thinks that water would help the poor chap, and falsely, though perhaps justifiably, believes that his pocket mat contains water. Consequently, curing the man from methanol poisoning cannot be the joint manifestation of Olos' and Ounas' competences, since Olos is incompetent. Yet, Olos' actions are partially responsible for Ounas' success, since she wouldn't have been able to cure the man had Olos not given the man ethanol. Therefore,

¹⁶ Molnar (2003), for instance, talks of joint manifestations of numerous dispositions (though not of dispositions of numerous different objects or agents).

according to the joint manifestation account, Ounas deserves less credit for the success if Olos is incompetent than if he is competent. But whether Olos is competent strikes us as completely irrelevant to how much credit Ounas deserves for the success. Similarly, her action is equally morally worthy irrespective of whether Olos is competent. So the joint manifestation account cannot be correct.

In the beginning of this section we flagged as problematic the way in which the Cake Theory treats the contributions of other agents to one's success (as in cases of testimony) on a par with environmental contributions that may - at least in many cases - plausibly be thought of as introducing an element of luck. But we think the problem with the Cake Theory runs deeper. The deeper problem is the way in which contributions of other factors are thought to always detract from the credit deserved by an agent. This is the core idea of the Cake Theory, as expressed by its second tenet:

- (2) The degree to which a success is creditable to an agent A is a function of the degree to which A's competence, as opposed to other factors, is responsible for the success.

Indeed, we don't think it is important that the extra amount of ethanol received by the poisoned man has its source in the intentional actions of another agent. Had the man ended up ingesting some extra ethanol due to some bizarre chain of events not connected with the actions of agents, Ounas would still have deserved credit for saving the man from methanol poisoning. In short, we think the Cake Theory identifies the wrong facts as those grounding degrees of credit.

7. CONCLUDING REMARKS

We've argued that extant virtue epistemological views are intricately connected to the Cake Theory of credit, and that the second tenet of the Cake Theory is false. Factors external to one's agency may be responsible for a success even to a significant degree without detracting from the credit one deserves for the success. We have given cases in which multiple agents each deserve a high degree of credit for a single success. But the Cake Theory, due to its structure, rules out such cases. Recall, for instance, the case *Methanol and Ethanol*, in which Olos and Ounas save a man from methanol poisoning. We don't think that Olos, for instance, deserves any less credit for saving the man in the case we described than in a different one in which he carries a slightly larger pocket mat, one big enough to contain just enough ethanol to save the man even without the help of Ounas.

But if the degree of credit an agent deserves for a success isn't a function of how responsible the agent's competences, as opposed to other factors, are for the success, then what is it a function of? Here we lack the space to develop a full-fledged theory. Instead we offer some desiderata that theories of credit ought to satisfy, and gesture in what we see as the right direction.

A theory of credit should:

- (i) make sense of the idea that credit can come in degrees,
- (iii) allow for credit to be shared in a way that does not detract from the degree of credit deserved by the involved parties, and,
- (iii) be applicable outside epistemology.

In the case of (iii) we have, in particular, in mind the common view on which morally worthy action is a matter of being creditable for doing the morally right thing. In addition to these three desiderata, many virtue epistemologists hold that we can give an account of knowledge in terms of credit: knowledge is creditable true belief, a sufficient amount of credit ruling out the kind of luck that is incompatible with knowledge. Though we think the notion of credit has important work to do in epistemology, we doubt that knowledge is creditable true belief. Indeed we think there are cases of knowledge that *don't* involve manifesting competences to believe truly -- or, indeed, competences to know.¹⁷

We think virtue epistemologists are right in attempting to ground facts about credit in properties of agents: for a success to be attributable to an agent, after all, it must be suitably connected to something pertaining to the agent. We don't, however, have to think of the relevant properties as *competences*, thought of as dispositions succeed. Instead, they might simply be the dispositions manifested by an agent, dispositions manifesting themselves as the relevant attempt (the forming of a belief, performing of an action, etc.). Further, we think the most promising way of thinking about the kind of luck or accidentality at issue, one incompatible with credit and attributability, is modal. Start with the rough and ready idea that an agent deserves credit for a success *S* (in a case *c*) if there is a modally robust connection within some relevant set of cases between manifesting the dispositions that the agent manifests (in *c*) and *S*. (Obviously the dispositions that enter the explanation here cannot be dispositions to *S*, since a disposition to *S* cannot have non-*S*-ings as its manifestations.) However, we need not think of the relevant cases as ones that are modally close.

One way to fill in the details, defended in Lasonen-Aarnio (forthcoming), takes as its starting point the contrast between being a fluke, coincidence, or accident, on the one hand, and having an explanation, on the other. A success that is attributable and hence, creditable, to an agent must depend in a specific way on the dispositions manifested by the agent. In particular, there must be an explanatory generalization connecting the dispositions and success, a generalization that explains the token success. The account is broadly modal, for explanatory generalizations must be sufficiently *invariant* across

¹⁷ See Lasonen-Aarnio (2010) and Hirvelä (2018) on knowing without credit for true belief, and Hirvelä (2019b) and Lasonen-Aarnio (2021, forthcoming) for knowing without having the competence to do so.

cases involving different *interventions* and background conditions (cf. Woodward (2003)).¹⁸ In sum, for a success *S* to be creditable to an agent, there must be a sufficiently modally robust, invariant connection between the dispositions manifested by the agent and *S*, a connection that explains the fact that *S* occurs on this occasion.¹⁹ On such a view, the degree of credit the subject deserves depends on how robust the generalization is, and/or how well it explains the token success.

Another way to cash out the rough and ready idea would be to maintain that credit is a function of how characteristic, or typical, *F*-ing, in the way in which *S* actually *F*'s, is as a manifestation of the dispositions that the agent manifests. On this view *F*-ing would be a typical or characteristic manifestation of a set of dispositions *D*, only if there is some normal world where the *D* manifests as *F*-ing when triggered.²⁰ The larger the set of normal worlds across which *D* manifests as *F*-ing when triggered the more typical or characteristic *F*-ing is as the manifestation of *D*.

We could then maintain that if the dispositions that the agent manifests in believing *p* typically manifest as knowledge, and *S* knows that *p*, then her knowing *p* would be creditable to her. An agent would deserve credit for doing what is morally right if the dispositions that they manifest in doing what is morally right normally would result in doing what is morally right in their circumstances. The degree of credit an agent deserves for *F*-ing would be a function of how large the set of normal worlds across which the dispositions that the agent manifests manifest as *F*-ing when triggered.

These are vague suggestions, but they do look promising with respect to the desiderata we laid down in that they might allow us to share our cake and have it too.

Acknowledgments: We would like to thank J. Adam Carter, Ninni Suni, and the 'Ethics, Epistemology, and Normativity' group at Helsinki for helpful comments and discussion. This project has received

¹⁸ We should be very cautious about drawing conclusions from successful causal explanations to claims involving quantitative causal notions, such as claims about the magnitude of the total effect of one variable on another, or claims about comparative causal relevance--indeed, Woodward (2003: 38) notes that his approach has nothing to say about such quantitative causal notions.

¹⁹ Various things can interfere with one's success, and in any case some cooperation from the environment (and perhaps other agents) will be required - hence, the relevant generalizations will be very far from fully invariant. But this is also true of a wide range of generalizations deployed in scientific explanations, ones standardly deployed in the social sciences being a case in point. And as Woodward (2003) likes to point out, even generalizations deployed in the natural sciences are not fully invariant: for instance, the field equations of general relativity break down when quantum gravitational effects become important.

²⁰ For use normal worlds see for instance M. Smith (2016) and Goodman and Salow (2018).

funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 758539.

REFERENCES

- Baehr, J. (2012). Credit Theories and the Value of Knowledge. *Philosophical Quarterly*, 62(246), 1-22. doi:10.1111/j.1467-9213.2011.698.x
- Bradford, G. (2015). Knowledge, Achievement, and Manifestation. *Erkenntnis*, 80(1), 97-116. doi:10.1007/s10670-014-9614-0
- Carter, J. A. (2016). Robust Virtue Epistemology As Anti-Luck Epistemology: A New Solution. *Pacific Philosophical Quarterly*, 97(1), 140-155. doi:10.1111/papq.12040
- Gaultier, B. (2014). Achievements, Safety and Environmental Epistemic Luck. *Dialectica*, 68(4), 477-497. doi:10.1111/1746-8361.12081
- Goodman, J., & Salow, B. (2018). Taking a chance on KK. *Philosophical Studies*, 175(1), 183-196. doi:10.1007/s11098-017-0861-1
- Greco, J. (2003). Knowledge as Credit for True Belief. In M. DePaul & L. Zagzebski (Eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Greco, J. (2010). *Achieving Knowledge: A Virtue Theoretic Account of Epistemic Normativity*. Cambridge: Cambridge University Press.
- Greco, J. (2016). Knowledge, Virtue, and Safety. In M. A. Fernández Vargas (Ed.), *Performance Epistemology - Foundations and Applications* (pp. 51-61). Oxford: Oxford University Press.
- Greco, J. (2018). Virtue, Knowledge, and Achievement. In *The Routledge Handbook of Virtue Epistemology*: Routledge.
- Hirvelä, J. (2018). On Virtue, Credit and Safety. *Grazer Philosophische Studien*, 95(1), 98-120.
- Hirvelä, J. (2019a). Global safety: how to deal with necessary truths. *Synthese*, 196(3), 1167-1186. doi:10.1007/s11229-017-1511-z
- Hirvelä, J. (2019b). Knowing Without Having The Competence to Do So. *Thought: A Journal of Philosophy*, 8(2), 110-118. doi:10.1002/tht3.411
- Hirvelä, J. (2020). No Safe Haven For The Virtuous. *Episteme*, 17(1), 48-63.
- Hirvelä, J., & Paterson, N. J. (2021). Need knowing and acting be SSS-Safe? *Thought: A Journal of Philosophy*, 10(2), 127-134. doi:<https://doi.org/10.1002/tht3.487>
- Kelp, C. (2017). Knowledge First Virtue Epistemology. In A. Carter, E. Gordon, & B. Jarvis (Eds.), *Knowledge First: Approaches in Epistemology and Mind* (pp. 224-245). Oxford: Oxford University Press.
- Lackey, J. (2007). Why We Don't Deserve Credit for Everything We Know. *Synthese*, 158(3), 345-361. doi:10.1007/s11229-006-9044-x
- Lasonen-Aarnio, M. (2010). Unreasonable Knowledge. *Philosophical Perspectives*, 24(1), 1-21.
- Lasonen-Aarnio, M. (2021). Dispositional Evaluations and Defeat. In J. Brown & M. Simion (Eds.), *Reasons, Justification and Defeat* (pp. 93-115). Oxford: Oxford University Press.
- Lasonen-Aarnio, M. (forthcoming). Perspectives and Good Dispositions. *Philosophy and Phenomenological Research*.
- Littlejohn, C. (2014). Fake Barns and False Dilemmas. *Episteme-a Journal of Individual and Social Epistemology*, 11(4), 369-389. doi:10.1017/epi.2014.24
- Miracchi, L. (2014). *Getting Things Done*. (Ph.D. Thesis), Rutgers University,
- Miracchi, L. (2015). Competence to Know. *Philosophical Studies*, 172, 29-56.
- Molnar, G. (2003). *Powers: A Study in Metaphysics: A Study in Metaphysics* (Vol. 72): Oxford University Press.
- Prior, E. W., Pargetter, R., & Jackson, F. (1982). Three theses about dispositions. *American Philosophical Quarterly*, 19(3), 251-257.
- Pritchard, D. (2005). *Epistemic Luck*. Oxford: Oxford University Press.
- Pritchard, D. (2010). Knowledge and Understanding. In D. Pritchard, A. Millar, & A. Haddock (Eds.), *The Nature and Value of Knowledge* (pp. 5-88). Oxford: Oxford University Press.
- Pritchard, D. (2012). Anti-Luck Virtue Epistemology. *Journal of Philosophy*, 109(3), 247-279.

- Riggs, W. (2002). Reliability and the Value of Knowledge. *Philosophy and Phenomenological Research*, 64, 79-96.
- Sliwa, P. (2016). Moral Worth and Moral Knowledge. *Philosophy and Phenomenological Research*, 93(2), 393-418.
- Smith, A. M. (2015). Responsibility as Answerability. *Inquiry: An Interdisciplinary Journal of Philosophy*, 58(2), 99-126.
- Smith, M. (2016). *Between Probability and Certainty - What Justifies Belief*. Oxford: Oxford University Press.
- Sober, E. (1988). Apportioning Causal Responsibility. *The Journal of Philosophy*, 85(6), 303-318. doi:10.2307/2026721
- Sosa, E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge: Volume I*. Oxford: Oxford University Press.
- Sosa, E. (2009). *Reflective Knowledge: Apt belief and Reflective Knowledge, Volume II*. Oxford: Oxford University Press.
- Sosa, E. (2015). *Judgment and Agency*. Oxford: Oxford University Press.
- Sosa, E. (forthcoming). *Epistemic Explanations*. Oxford: Oxford University Press.
- Turri, J. (2011). Manifest Failure: The Gettier Problem Solved. *Philosophers Imprint*, 11(8), 1-11.
- Watson, G. (1996). Two Faces of Responsibility. *Philosophical Topics*, 24(2), 227-248.
- Vetter, B. (2016). Are abilities dispositions? *Synthese*. doi:10.1007/s11229-016-1152-7
- Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*: Oxford University Press.
- Zagzebski, L. (1994). The Inescapability of Gettier Problems. *Philosophical Quarterly*, 44(174), 65-73.