

Knowledge without credit, exhibit 4: extended cognition

Krist Vaesen

Received: 19 June 2009 / Accepted: 25 March 2010 / Published online: 8 April 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The Credit Theory of Knowledge (CTK)—as expressed by such figures as John Greco, Wayne Riggs, and Ernest Sosa—holds that knowing that p implies deserving epistemic credit for truly believing that p . Opponents have presented three sorts of counterexamples to CTK: S might know that p without deserving credit in cases of (1) innate knowledge (Lackey, Kvanvig); (2) testimonial knowledge (Lackey); or (3) perceptual knowledge (Pritchard). The arguments of Lackey, Kvanvig and Pritchard, however, are effective only in so far as one is willing to accept a set of controversial background assumptions (for instance, that innate knowledge exists or that doxastic voluntarism is wrong). In this paper I mount a fourth argument against CTK, that doesn't rest on any such controversial premise, and therefore should convince a much wider audience. In particular, I show that in cases of extended cognition (very broadly conceived), the most salient feature explaining S 's believing the truth regarding p may well be external to S , that is, it might be a feature of S 's (non-human, artifactual) environment. If so, the cognitive achievement of knowing that p is not (or only marginally) creditable to S , and hence, CTK is false.

Keywords Knowledge · Credit · Extended cognition

1 Introduction

Consider a lousy archer, called Franz. Franz places a lousy shot, but, due to the fortunate interference of a strong western wind, hits the boar nonetheless. The boar is brought down through luck, rather than through Franz' skill, so although the shot is attributable to him, the success is not.

K. Vaesen (✉)
Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: k.vaesen@tue.nl

This is a fairly typical way of framing what is special about knowledge. Knowledge is like hitting the boar through skill rather than through luck; it is believing the truth *because of* the correct application of one's own cognitive abilities. So we would not attribute knowledge to someone in case she holds a true belief that p , but her truly believing that p is explained by, say, the environment accidentally making her belief true. Stated otherwise: to know that p is to deserve credit for believing the truth regarding p —voilà, the Credit Theory of Knowledge (CTK) in its most basic form. One can find it in, among other places, Greco (2007),¹ Riggs (2007)² and Sosa (2007).³

The strategy of those who oppose CTK is straightforward: find an instance of knowledge in which the credit condition isn't met. Three such counterexamples have been suggested. Both Lackey (2007) and Kvanvig (2009) argue that *innate knowledge* (at least: *if* it exists) would effectively undermine CTK—and indeed, *prima facie* it seems pretty awkward to credit someone for the innate beliefs she happens (or better: *would* happen) to have. Second, there is *testimonial knowledge* (see Lackey 2007, 2009). Lackey argues that one can acquire knowledge by, so to speak, passively absorbing what one is told; if so, credit should go, if anywhere, to the testifier rather than to her audience. And third, Pritchard (2005a) has expressed his doubts regarding credit theories in simple cases of *perceptual knowledge*: since basic perception is not within one's immediate control, speaking of creditworthiness seems problematic at least.

Each of these counterarguments has its own problems, though. First, the argument of *innate knowledge* obviously is compelling in so far as (one believes that) such knowledge exists—a controversial conjecture, to say the least. Moreover, even if one is willing to concede the existence of innate knowledge, the CTK-theorist can easily downplay its significance; Riggs (2009), for example, says to have no problem “limiting” his brand of CTK to the most common cases of knowledge, namely those involving empirical (as opposed to innate) knowledge.

The same author also questions the effectiveness of Lackey's case of *testimonial knowledge*. In Lackey's scenario, a person called Morris, who has just arrived at Chicago train station, forms a true belief about the whereabouts of the Sears Tower, after having asked the first adult passerby. Lackey argues that while Morris knows the location of the Sears Tower, he doesn't deserve epistemic credit for his believing the truth; hence, knowledge and credit do not always go hand-in-hand. Riggs, now, simply doubts that Morris has knowledge regarding the Sears Tower (*ibid.*, p. 11): ‘Why on earth would we say that Morris knows where the tower is when he has picked a stranger at random, and unhesitatingly (and, one assumes, unreflectively) accepted what that person said? On the face of it, this is terrible epistemic practice.’

The dilemma Lackey faces is this. On one hand, she may raise the standards for testimonial knowledge. She could assume that Morris doesn't simply open up his brain, but does some extra cognitive work; Morris, for instance, might select not the first passerby, but walk to a tourist office and ask for information there. If so, the

¹ Greco (p. 57): “[...] knowledge attributions can be understood as credit attributions: when we say that someone knows something, we credit them for getting it right.”

² Riggs (p. 329): “[...] knowledge is, complications aside, credit-worthy true believing.”

³ Sosa (p. 92): “[b]elief amounts to knowledge when apt: that is to say, when its correctness is attributable to a competence exercised in appropriate conditions.”

CTK-theorist presumably accepts that Morris through testimony comes to *know* about the Sears Tower; the problem now is, however, that Morris' discriminatory behavior (i.e. carefully selecting a reliable testimonial source) makes him deserve credit again.

Consequently, the effectiveness of Lackey's case against CTK strongly depends on how high one sets the standards for knowledge. My intuitions, for what it's worth, tend more toward Morris indeed acquiring knowledge; but for a strong, generic argument against CTK, a scenario indisputably involving knowledge would be helpful.

Third, according to Pritchard's *perceptual knowledge* argument, our lack of control over our perceptual (and many other) beliefs is at odds with us deserving credit for them. We seem to get such ideas "willy nilly", they are hardly *our* cognitive achievement. In light of this, the argument goes, CTK-theorists either should admit that perceptual knowledge is illusory, or give up their credit condition.

Unfortunately, as Riggs (2007) notes, quite a number of philosophers have argued a case for exactly the opposite; they contend that people *do* have sufficient control over their beliefs to be rightly attributed credit.⁴ These defenders of what is called "doxastic voluntarism" will not be impressed by Pritchard's charges. So unless the dispute between voluntarists and anti-voluntarists is settled first, the prospects for a refutation of CTK that is palatable for all (or most) look dim.

In sum, what Kvanvig's, Lackey's, and Pritchard's counterarguments have in common is that they can perform as intended, only when some extra, contested assumptions are made—respectively, about the existence of innate knowledge, about the standards for knowledge, and about our lack of doxastic control. And that is, obviously, their weakness.

In this paper I develop a fourth line of argument against CTK, the efficacy of which, I think, isn't conditional on any such contestable premise(s). It concerns knowledge acquired through extended forms of cognition. Those familiar with recent discussions in the philosophy of mind might at this point dismiss my argument already, for to many the hypothesis of extended cognition is contestable *par excellence*. But to reassure the reader: the brand of extended cognition my argument relies on is, as will become apparent, quite weak, hence easily digestible. Basically, the only thing one needs to accept is that humans may use cognitive aids to produce cognitive outputs, that we may acquire knowledge by putting to work simple things like glasses, thermometers and computers. Indeed, my argument works whenever one is willing to subscribe to this quite trivial claim.

The baseline of my argument then will be this: *S* can come to know that *p*, even when the most salient feature explaining *S*'s true belief is a feature of *S*'s (cognitive) artifactual environment. But whereas in Gettier-style cases the environment's contribution is a fortunate coincidence, it is intended (non-lucky) in the scenarios I have in mind.

As a rough-and-ready preview to my argument, let me draw a parallel to archery again. Suppose Franz, our lousy archer, hits a target, not through skill, but thanks to a bow-and-arrow set, neatly installed on a tripod, perfectly aimed at the target,

⁴ Riggs refers to Hieronymi (2006), Raz (1999), Adler (2002), Owens (2000), and Audi (2001).

and equipped with a simple “shoot”-button. Knowledge might be like this; hitting the target, while all relevant cognitive work was delegated to external aids.

2 What epistemic credit is (supposed to be)

Let's sort out some terminological issues first. What exactly does the CTK stand for? What does it mean to say that knowledge is credit-worthy true belief? In answering these questions, I will restrict myself to two authors, John Greco and Wayne Riggs, since they both are quite explicit on this score. My case will be particularly strong if I succeed in refuting both CTK_{Greco} and CTK_{Riggs}.

Take CTK_{Greco} first. According to Greco (2003), *S* knows that *p*, iff *S* deserves intellectual credit for believing the truth regarding *p*. Intellectual credit, in turn, requires that:

- (G.1) believing the truth that *p* has intellectual value;
- (G.2) believing the truth regarding *p* can be ascribed to *S*;
- (G.3) *S*'s reliable cognitive character is an important necessary part of the total set of causal factors that give rise to *S*'s believing the truth that *p*.

Condition (G.3) is crucial, since it is here that cognitive skill is introduced so as to separate the wheat (knowledge) from the lucky chaff (accidentally true belief). Lackey (2007) points out, though, that Greco doesn't endorse so much (G.3) as (G.3*):

- (G.3*) *S*'s reliable cognitive character is the *most salient* part of the total set of causal factors that give rise to *S*'s believing the truth that *p*.

The reason for that is simple. In Gettier-type cases, such as Chisholm's dog-looking-like-a-sheep scenario,⁵ a necessary part of the total set of causal factors giving rise to *S*'s truly believing that *p* still *is* part of *S*'s cognitive character, namely: *S* relying on her perceptual faculties to form the belief that *p*. To deny *S* knowledge in Gettier-style scenarios, Greco himself writes that *S*'s cognitive character 'is [in such cases] not the *most salient* part [of the total set of causal factors that give rise to *S*'s believing the truth that *p*] (Greco 2003, p. 130, italics mine)'. Hence the plausibility of (G.3*).

Lackey raises another interesting issue. How should we understand cognitive character in (G.3*)? Does it simply refer to the cognitive faculties *S* *actually* applied to form the belief that *p* or is it something more stable, a trait of *S* which is *typically* revealed when she is cognizing? In the latter case, it is particularly hard for a lousy thinker to come to know something extraordinary, that is, something unexpected given her stable (and lousy) cognitive character. Therefore, Lackey suggests the following amendment:

- (G.3**) *S*'s reliable cognitive *faculties* are the most salient part of the total set of causal factors that give rise to *S*'s believing the truth that *p*.

⁵ The scenario is as follows. *S* forms the belief that there is a sheep in the field, based on her seeing a sheep-looking dog. *S* belief happens to be true, though. There actually *is* a sheep in the field, invisible to *S*, grazing behind a rock.

What (G.3**) is supposed to do, thus, is to get into focus the *actual* cognitive faculties applied by *S*; it is these that should be decisive in (withholding) attributions of credit, not *S*'s track record of cognitive achievements. So if a lousy thinker would “reliably” stumble over a brilliant discovery, she would deserve credit (and thus the stamp of knowledge), even though the discovery wouldn't be representative for her typical cognitive bent.

In the remainder, whenever referring to CTK_{Greco} , I will assume (G.3**) rather than (G.3) or (G.3*), since (G.3**) seems the most convincing interpretation available. Lowering the stakes—that is adopting (G.3) or (G.3*)—would, I suppose, make the rest of my argument less compelling.

Then there is CTK_{Riggs} —which, according to the author (Riggs 2009), differs in important respects from CTK_{Greco} . CTK_{Riggs} is basically an “anti-luck” theory of knowledge. For Riggs, saying that something is due to luck means it is not attributable to the beneficiary (or victim) of the luck in question. Credit, as a cognate of attributability, then ‘is simply a shorthand for saying that some event, state of affairs, or consequence thereof is attributable to an agent, as an agent (ibid., p. 3).’ Applied to matters of knowledge: *S* knows that *p* only if *S*'s holding the true belief that *p* is attributable to her as a cognitive agent.

But to what kinds of luck is attributability supposed to be antithetical? Riggs discerns two types. First, there is the veritic luck of Gettier-style cases (such as Chisholm's dog-looking-like-a-sheep scenario, see footnote 5). In such cases, *S*'s coming to hold a true belief that *p* is not ‘the product of *S*'s actual [cognitive] abilities (Riggs 2007, p. 335)’, and hence, *S* does not really *know* that *p*. Second, attributability as anti-luck implies that *S* coming to hold a true belief that *p* wasn't accidental, in the sense of being (sufficiently) caused by *S*'s intention to have true beliefs. Riggs (1998) gives the following example. Suppose I have patent evidence that my business partner is guilty of fraud. Yet, because I have deep inhibitions against becoming involved in conflicts, I go to great mental lengths to explain away the evidence and to believe my partner is innocent. In fact, my partner really *is* innocent; he is just being framed by someone else. So while my belief about my partner is true, it doesn't qualify as knowledge, since it is true *accidentally*. My actions aren't guided by my intention (or better: desire) to have true beliefs any longer, but by another desire (*viz.* the desire to avoid confrontation). In such cases, my cognitive success is, as Riggs puts it, inadvertent.

Putting this together, Riggs thinks that (ibid., p. 335): *S* knows that *p* iff *S* holds the true belief that *p* and:

- (R.1) *S*'s coming to hold a true belief in this instance is the product of *S*'s actual abilities; and
- (R.2) *S*'s coming to hold a true belief in this instance is not inadvertent.

Given this formulation, the difference between CTK_{Greco} and CTK_{Riggs} appears pretty small. After all, (R.1) is a statement concerning causal efficacy: we want ‘the outcome [i.e. *S* coming to hold the true belief that *p*] [to be] causally due to the agent's abilities (ibid., p. 335)’. Or more in line with Greco's terminology: the most salient causal feature explaining the true belief must be *S*'s cognitive faculties. The

fact that Riggs talks about *actual* abilities makes (R.1) indistinguishable perhaps not from (G.3) or (G.3*), but quite plausibly from (G.3**).⁶

Regarding (R.2): indeed, such a claim is absent in CTK_{Greco} , but two points in response. First of all, as Riggs admits, it is quite hard to articulate what “inadvertently coming to form a true belief that p ” actually amounts to. Second, even if one could spell out condition (R.2) in sufficient detail, I can allow myself to focus on condition (R.1) alone; given the biconditional (“iff”), CTK_{Riggs} is false just in case one of its conditions doesn’t hold. This is precisely what I will show in subsequent sections.

To recap, CTK_{Greco} and CTK_{Riggs} share one intuition, that can be formulated either as (G.3**) or as (R.1). As such, it suffices to rebut that shared intuition in order to rebut both CTK_{Greco} and CTK_{Riggs} . Since Greco was the first to propose a theory of knowledge in terms of credit-worthy belief and for purposes of simplicity, I will adopt Greco’s terminology and work only with (G.3**). One should bear in mind, however, that my criticism applies just as much to (R.1), and thus to CTK_{Riggs} .

3 Credit and extended cognition: first pass

In this section and the next, I introduce two examples of what one may call extended cognition (e-cog), a thesis proposed fairly recently in the philosophy of mind. According to proponents of e-cog (Andy Clark, Mark Rowlands, John Haugeland, to name but a few), cognitive processes may take place outside the boundaries of the human skin—*extending* into the subject’s environment—implying that the human skin doesn’t delimit the thinking subject. One prototypical example (given by [Clark and Chalmers 1998](#)) is Otto, an Alzheimer patient who uses a notebook as a substitute for his failing memory. Otto relies on his notebook in the same way as a healthy person relies on her internal memory; given this functional equivalence between internal (biological) memory and external (artifactual) memory, Clark and Chalmers claim that there is no principled reason to consider the former as a cognitive entity, but not the latter:

If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is [...] part of the cognitive process. ([Clark and Chalmers 1998](#), p. 8)

Much has been said about the plausibility of e-cog. One obvious problem is conceptual: despite functional equivalence, external processes do not seem to meet traditional criteria of the cognitive. [Adams and Aizawa \(2001\)](#), for instance, argue that real cognitive processes involve non-derived content (which e-cog processes don’t), and that the causal mechanisms underlying external and internal processes are too different to form a cognitive kind.⁷

⁶ Put otherwise, I don’t think that the difference between Greco and Riggs is a salience requirement (as Riggs suggests, p. 1); in my opinion, causal salience and efficacy can be construed as equivalent. The difference rather is (i) in *typical* (Greco) versus *actual* (Riggs) applications of cognitive abilities as a litmus test for credit, and (ii) in consideration (R.2).

⁷ For another much cited critique of e-cog, see [Rupert \(2004\)](#).

These are thorny issues indeed,⁸ but whatever we decide the mark of the cognitive to be, e-cog contains a fairly uncontroversial (but in epistemology manifestly underplayed)⁹ part: the fact that human cognition is strongly dependent on external resources (whether or not we call them cognitive). Some features of the world actively scaffold us in our cognitive endeavors and as such are causally relevant to the kinds of beliefs we happen to have. And as long as CTK-theorists recognize this, my argument will appear effective. The ultralight version of e-cog I will exploit, thus, is supposed to be attractive to a wide audience and sidesteps the conceptual morass surrounding the notion “cognitive”.

A very basic example of external scaffolding is in fact invoked quite commonly in epistemology (but not recognized as such): *S* using a thermometer to form beliefs about her environment. In matters of temperature, *S* indeed is quite dependent on a device like a thermometer.¹⁰ So an external feature of *S*’s external environment (i.e. the thermometer) certainly *is* causally relevant to *S*’s true beliefs regarding, say, room temperature. And the instrument is not one of those environmental features (prominent in Gettier-type cases) that contribute only accidentally to true belief; no, the device is *designed* and *used* to play the particular role it plays.

Let me point out how this might bear on CTK.¹¹ Consider the following scenario, due to Pritchard (2005b). Ferdinand likes to be up to date with respect to room temperature, and therefore regularly consults a thermometer on the wall. Ferdinand doesn’t know the thermometer is defective and that it fluctuates randomly within a given range. Nonetheless, Ferdinand’s readings are reliable, since unbeknownst to him, a benevolent demon sees to it that, whenever Ferdinand consults the thermometer, room temperature is adjusted so that it actually corresponds to what is displayed on the device. Although reliably formed, Ferdinand’s beliefs do not count as knowledge. Why not? According to CTK: Ferdinand doesn’t know, because his cognitive faculties are not the most salient feature explaining his correct temperature readings. He doesn’t deserve credit for his true beliefs; if credit (as causal salience) should go to anyone, it would be to the benevolent demon.

But what would happen if we removed the benevolent demon and let Ferdinand use a perfectly functioning thermometer? He would figure out the temperature correctly, of course. But would his cognitive faculties really be the most salient feature explaining his success? As he is helped in the demon case (viz. by the demon), he now seems helped just by another agent: the manufacturer of the device.¹² If not for her, Ferdinand wouldn’t have temperature beliefs at all. If the manufacturer doesn’t at least earn some epistemic credit, a striking (and I guess unfair) inequality arises: if

⁸ This is not to say that they can’t be handled satisfactorily. Quite convincing e-cog (counter)arguments are e.g., Clark (2007), Menary (2006) and Rowlands (2009).

⁹ See Pritchard (forthcoming, p. 2), for a similar remark.

¹⁰ This in fact holds for all sorts of instrumentation; many augment (e.g. telescopes) or substitute (e.g. calculators) some of the cognitive faculties we are naturally endowed with.

¹¹ Recall that CTK here (and in the remainder) assumes (G.3**).

¹² Not to mention the efforts of those scientists involved in the invention of temperature. For an overview of this intricate achievement, one may read Hasok Chang’s *Inventing temperature: Measurement and scientific progress* (2004).

something goes wrong, we most likely blame the manufacturer, but if all goes well, we leave her out of the credit equation.

The simple point I am trying to make is in fact structurally similar to Lackey's argument regarding innate knowledge (Lackey 2007). As said, Lackey contends that the most salient explanatory feature for an innate true belief wouldn't be some cognitive trait of the person entertaining it, but the origin of the belief, such as natural selection or some other evolutionary mechanism. What this suggests is that outcomes such as true belief may be generated by causal factors extending across the boundaries of the individual cognizer. In Lackey's case, though, these external factors are (would be) outside human control, undeserving attributions of credit, whereas in the thermometer scenario, the causal contributions are within the realm of human intent—and therefore a proper target for normative appraisal. Another strength of the thermometer example is, of course, that it doesn't invoke a dubitable entity like innate knowledge.

Nonetheless, the scenario is vulnerable to at least two objections (more will follow in Sect. 5). The first one is that causal relevance isn't the same as causal salience. So the CTK proponent might concede that several external factors contribute to Ferdinand's success, but stress that *salience* is in the truth-conducive cognitive work performed by Ferdinand, not that performed by the manufacturer (let alone by the thermometer itself). It is Ferdinand who has decided to use the thermometer in the first place; it is Ferdinand who has developed a sense of trust regarding the reliability of the instrument, either through inductive generalization (trial-and-error, perhaps) or through testimonial evidence (provided by the manufacturer, who Ferdinand has chosen to trust).¹³ Ferdinand (no one else) is in control of the thermometer, so it makes good sense to consider him as causally privileged.

Second, Greco and Riggs might point out that causal salience is not supposed to explain *S* merely believing that *p*, but, as stated in (G.3**), *S* believing the *truth* about *p*. To demonstrate the causal salience of an external feature, therefore, I shouldn't refer to a counterfactual in which no belief is produced—such as Ferdinand not forming temperature beliefs in the absence of a thermometer—but to one in which, given similar cognitive processes, a false (rather than a true) belief is formed.

I think both objections are pertinent. So instead of trying to counter them, I will rather translate them into requirements for the design of a more effective scenario. As a refutation of CTK, such a scenario should thus at least: (i) downplay the cognizer's control in his coming to know that *p*—instead of *S* pulling information out of the device, the device should as it were push truths toward *S*; and (ii) involve external features explaining *true* belief, not just (the presence of) belief. I give it another try below.

4 Credit and extended cognition: second pass

The e-cog scenario developed in this section takes its cue from a book titled *The human factor* (2003) by cognitive engineer Kim Vicente, and concerns airport security, in particular the X-ray scanners used for baggage inspection. Such scanners make

¹³ This point is made for instance in Sosa (2006, p. 118).

the life of inspectors easier in a way (no need to open up and check the inside of every piece of luggage), but also more tedious and demanding. After all, illegal objects are exceedingly rare, which implies that operators typically have to stare for long uninterrupted periods of time at a monitor showing unsuspecting (and boring) goods (like toothbrushes, deodorants, computers, books). Under such conditions, people's vigilance decreases dramatically; after half an hour, performance drops by almost 50% (ibid., p. 138).

The solution (implemented worldwide after 9/11) relates to the fact that people remain more attentive if "false signals" are introduced periodically. In case of baggage scanners, these false signals are images of illegal objects superimposed on the actual images of the suitcases projected on screen. The false positives wake up the inspectors, and stimulate them to remain alert.

Suppose, now, that SYSTEM₁ is an ordinary pre-9/11 baggage scanner, whereas SYSTEM₂ is a post-9/11 upgrade, including a "false signal" engine. Whenever SYSTEM₂ projects a false image and the operator notices (she informs the system by, say, clicking on the image), the following message pops up: "False alarm: you were being tested!" If no message appears, the operator knows the threat is real. Consider, then, the following scenario:

SISSICASE: Sissi has been a baggage inspector all her life. She used to work with an old-fashioned SYSTEM₁, but since 9/11, the airport she is working for introduced a SYSTEM₂. Her supervisor Joseph, a cognitive engineer who was actually involved in the design of the device, has informed her how it works (how its operation is almost identical to the operation of the old system). Currently Sissi is inspecting a piece of luggage which contains a bomb. She notices and forms a true belief regarding the contents of the suitcase. As such, the bomb is intercepted and a catastrophe prevented from happening.

What is the most salient feature explaining Sissi's true belief? Note here that the relevant counterfactual is Sissi using the old-fashioned SYSTEM₁, resulting, we may reasonably suppose, in her failing to notice the bomb, and thus in her forming a *false* belief (rather than her just failing to form a belief). As such, requirement (ii), formulated at the end of the previous section, is met. Since Sissi's cognitive faculties are assumed constant over SISSICASE and the counterfactual, the difference between false and true belief is explained by a difference in the set-up of the machinery in both cases; hence, the most salient causal feature to the effect of true belief is external to Sissi.

The control condition (requirement (i)) is handled adequately, too. Salience is not in some discriminatory faculty of Sissi, deciding to replace SYSTEM₁ with a more truth-conducive SYSTEM₂, for that decision was Joseph's, her supervisor. Nor can it be in her treating Joseph as a reliable testimonial source, because Joseph plays an identical role in SISSICASE and its counterpart. We have no reason to suppose that Sissi applied different faculties of trust in either case; as an employee she trusts her supervisor, irrespective of whether he is explaining the workings of SYSTEM₁ or of SYSTEM₂.

My diagnosis, then, is this: Sissi *knows* about the bomb, but doesn't really deserve credit (in the sense of (G.3**)), simply because it is a feature external to her that makes

the difference between her believing the truth (“there is a bomb”) and her believing a falsehood (“no bomb in here”).

5 Objections and replies

In this section I will go through some objections one may raise.

Objection :: Partial credit One line of argument might go as follows. SISSICASE assumes that credit for true belief cannot be shared. CTK isn’t wedded to such an idea, though: a knower doesn’t need to get *all* credit for the knowledge she acquires. So perhaps Joseph (Sissi’s supervisor) or even the machine deserve *some* credit for their contribution to Sissi’s true belief, still the accomplishment is hers, and that suffices for the CTK-kite to fly.¹⁴

Reply First of all, this response goes quite some way in acknowledging Sissi’s strong dependence on external features (Joseph, the machinery)—a feather in e-cog’s cap, I would say. The problem is, however, that such a distribution of credit makes CTK vulnerable to Gettier-style charges again. For what the partial credit objection suggests is to downplay Sissi’s contribution from being the *most salient* feature to being just one feature among many; in other words, it is proposing (G.3) instead of (G.3*), whereas in fact, as I explained in Sect. 2, we need to make the opposite move, from (G.3) to (G.3*),¹⁵ and that to make CTK Gettier-proof. If Sissi forms the belief that there is a sheep in the field and her belief is true, not because what she sees is a sheep (it’s in fact a dog), but because a sheep, well-hidden behind a rock, *is* grazing in the field, Sissi’s cognitive faculties are relevant, but not salient enough to explain her true belief. Sissi’s true belief is a prototypical instance of *uncreditworthy* true belief; to accommodate for that, CTK theorists cannot content themselves with relevance, they should go for salience.¹⁶ Closing the circle: distributing credit—that is, acknowledging the many (not all salient) causal contributions to true belief—is attractive and natural, but not a cure we should expect to save CTK.

At this point CTK-theorists may deny that they have forgone the salience requirement. Distribution, they may argue, doesn’t mean equal distribution. Of all causal factors affecting Sissi’s belief, Sissi’s cognitive faculties remain privileged: her knowing about the bomb is *her* accomplishment, it is something *she* has done. This is what Riggs (2009) is after, I suspect, when introducing the notion of attributability: *S*’s true belief shouldn’t be attributable to luck, but to *S*.

¹⁴ Some such objection has been raised by Greco (2007, p. 65) with respect to Lackey’s case of testimonial knowledge. The partial credit objection expressed above is in fact a simple transposition of Greco’s and Riggs’ arguments to my case of extended cognition.

¹⁵ And ultimately, from (G.3*) to (G.3**), but let us ignore that for a moment.

¹⁶ Alternatively, as an anonymous reviewer pointed out, CTK theorists might simply dispute that in Gettier cases *S*’s cognitive faculties are a salient part of the causal story to begin with. If so, (G.3) remains valid, and consequently, my argument misses its target. Although this is possible, I find it hard to imagine how one can not just minimize, but just completely leave out Sissi’s faculties in a full causal explanation of her true belief. The mere fact that Sissi sees objects in front of her, does have at least some causal bearing on the truth of her judgments concerning these objects, doesn’t it? Because if not for her seeing these objects, there wouldn’t be not only no true judgments, but no judgments to begin with.

Now, let me say what I find unobjectionable: Sissi's belief formation is attributable to her, indeed. Her forming a *true* belief (and that's what at stake here), however, is *not*. Sissi's success is attributable to the machinery, I repeat. For suppose we were asked to explain the superior reliability of post-9/11 security systems, that is, why these yield more true beliefs regarding the contents of suitcases and the like. Here is an answer that wouldn't do: after 9/11, baggage inspectors simply got more attentive. This explanation misses the point, since the real cause for improvement wasn't a change in the cognitive character of the inspectors—inspectors didn't become more vigilant at once through their own force of will—but a change in the set-up of the artifactual environment.

The crucial truth-conducive work therefore has been done by Joseph, Sissi's supervisor. He found a way to bypass an element that was out of Sissi's control, namely her being liable to vigilance decrement, and as such increased the likelihood of Sissi forming true beliefs.

Objection :: Cognitive inequivalence Opponents might demur: your argument rests on the assumption that SISSICASE and its supposed relevant counterfactual are cognitively on a par. This is incorrect. There remains a significant difference between Sissi operating SYSTEM₁, versus her operating SYSTEM₂. For one thing, she has acquired some SYSTEM₂-specific skills, procedural knowledge about how to operate the new scanning device. In case she detects a suspicious object, for instance, she has learned to suspend judgment, and wait for a confirmation message to pop up (or not). Given these differences, her operating SYSTEM₁ is not the right counterfactual to SISSICASE; therefore, you have not demonstrated the causal salience of external scaffolds.

Reply SISSICASE and the counterfactual indeed display differences other than in the machinery. But I simply doubt that these differences are the most salient with respect to true belief. To repeat my question: what explains the superiority of post-9/11 systems? Wouldn't it be awkward to answer that higher reliability is due to operators having developed more truth-conducive skills? That it was a matter of them being better trained, of having learned to suspend judgment? There is at least one person who would object: Joseph.¹⁷ And the entire (cognitive) design community, for that matter.

To be clear, I do *not* deny that Sissi's efforts will figure in a full explanation of her true (rather than false) belief. I just think that they are not relevant enough to call them causally salient.

Objection :: No knowledge Another possible objection: SISSICASE is not a case of knowledge. When Sissi notices a suspicious object on-screen, and the scanner does *not* return a feedback message, she might form a belief on the contents of the suitcase in question, but this is not a full belief yet. Only when she opens up the piece of luggage and sees the bomb for herself, she can be said to fully believe, and thus to *know* about the bomb.

¹⁷ On the other hand, in cases of failure (Sissi not finding a bomb, causing a disaster), Joseph would arguably agree with such an arrangement, for failure wouldn't be attributable to him either; it would be Sissi's fault, her not having trained enough.

Reply If it is a matter of SISSICASE mistakenly assuming X-ray technology to allow for unambiguous bomb identification, we can easily adjust the scenario: Sissi's task could simply be to discern whatever it is that X-ray scanners are able to pick out unequivocally (human bones, perhaps?). If so, indications on-screen are fully reliable, and Sissi can have a full belief prior to opening the suitcase.¹⁸

One could of course doubt that indirect (i.e. instrument-mediated) perception can ever result in full belief. If one goes that way, one should also be skeptical about Ferdinand's temperature beliefs (see Sect. 3), and about the greater part of science. I will not even try to refute that sort of skepticism here, in part because there is no indication that CTK-theorists commit themselves to such a view.

Objection :: No knowledge (part 2) The "no knowledge" objection might take yet another form. One might argue that knowledge attributions are dependent on the purposes of the attributor. Greco (2008), for one, explicitly endorses such a position. If so, one can reasonably wonder why we should accept that SISSICASE is a case of knowledge, for whether Sissi knows, is just dependent on the interests and purposes of the person judging Sissi's performance.¹⁹

Reply This might be so, but to find out what caused the disparity in the truth of Sissi's beliefs, we must work with a *ceteris paribus* clause. That is, to figure out the effect of the machinery, we vary the machinery (SYSTEM₁ vs. SYSTEM₂), while keeping *all* other things equal, including the purposes of the person judging Sissi's epistemic achievements. So if the purposes of an attributor would make her attribute knowledge in case of SYSTEM₁, the same should hold for the case involving SYSTEM₂ (i.e. SISSICASE), and vice versa. In other words, attributor contextualism *sensu* Greco might be attractive, but it is a view orthogonal to my argument.

Objection :: E-cog is too weak Likely the most audacious counterargument is to object to the particularly weak reading of e-cog I have been endorsing. Why not agree with the extended mind theorists' strong reading of e-cog and consider the processes going on in the machinery as genuinely cognitive, as belonging to "Sissi the extended cognitive agent"? On this condition, Sissi's (now extended) faculties remain the most salient feature explaining her true belief, and CTK is saved.

Reply This might indeed be the dilemma: to adopt the strong programme of e-cog, or to abandon CTK. I leave it to others to argue for the first option, and give some more comments on the second in the section below.

6 Prospects for credit

Is there really no way, in the face of e-cog, to save the intuition that knowledge is (must be!) produced through cognitive skill? There is a way, I think, but it requires quite some tinkering.

¹⁸ Of course, if there were such a perfect, 1-to-1 correlation (say, between the presence of human bones and measured reflected radiation) Sissi wouldn't be needed for bone detection any more. But we could still ask her, for philosophical purposes, to double-check the machine.

¹⁹ I thank an anonymous reviewer for pointing out this possible objection.

Let me reconstruct how CTK was originally defined (making some simplifications for purposes of clarity):

(CTK) *S* knows that *p*, iff *S*'s reliable cognitive faculties are the most salient part of the total set of causal factors that give rise to *S*'s believing the truth that *p*.

SISSICASE undermines CTK's salience requirement, I argued, so it is useful to make an adjustment there. We can lower CTK's stakes, by requiring knowledge to involve *S*'s cognitive faculties in relevant (rather than salient) ways:

(CTK*) *S* knows that *p*, iff *S*'s reliable cognitive faculties are an important part of the total set of causal factors that give rise to *S*'s believing the truth that *p*.

(CTK*) already sits a bit more comfortable with SISSICASE. Indeed, the machinery isn't the sole cause of Sissi's true belief; Sissi's perceptual skills, for instance, are relevant, though not the difference-maker yielding true belief. To make (CTK*) Gettier-proof, though, a second adjustment is needed. For suppose Sissi forms the belief that there is a bomb in the suitcase and this is true, not because what she sees is a bomb (it is in fact a hairdryer), but because behind the hairdryer a bomb is hidden. In this case, Sissi's cognitive faculties are important to her true belief, yet this doesn't make for knowledge. Hence, we should rewrite (CTK*) as follows:

(CTK**) *If S* knows that *p*, *then S*'s reliable cognitive faculties are an important part of the total set of causal factors that give rise to *S*'s believing the truth that *p*.

(CTK**) relaxes the conditions for knowledge a bit further. Instead of using cognitive faculties to define what knowledge is, cognitive faculties are considered a feature present just whenever knowledge is. And if (CTK*), save Gettier, accommodated for e-cog, (CTK**) will, by extension, do likewise.²⁰

In sum, (CTK**) is plausible, but limited in scope. It points to an important characteristic of knowledge—knowledge implies credit, somehow, somewhere—but doesn't pick out a unique property discriminating it from lucky true belief—the problem CTK-theorists set out to resolve. So although informative, (CTK**) will likely be a disappointment for those seeking the truth about knowledge.

7 Conclusion

Suppose again that Franz, our lousy archer, pushes a button on a remote control, thereby activating a bow-and-arrow set, perfectly positioned and calibrated by a more capable archer. Suppose moreover that the shot hits target. Although the shot might be attributable to Franz, hitting target isn't.

²⁰ In a recent paper—to my mind, the only manuscript exploring the epistemological ramifications of e-cog in detail—Pritchard (forthcoming) too suggests that only a weak version of cognitive ability (à la (CTK**)) can be made consistent with e-cog. In his argument, though, he dismisses stronger versions (à la (CTK)) not on grounds particular to e-cog, but by rehearsing Lackey's case of testimonial knowledge. Only in a second step he discusses e-cog, showing its fit with weak theories of credit. I have shown in this paper, however, that we can make Pritchard's *first* step independent from Lackey; indeed, I introduced e-cog as an alternative (or complement) to her case of testimony.

The same goes, I have argued, for knowledge: S may come to know that p , not through S 's own skill, but due to S 's environment being intentionally set-up such that S would come to believe the truth regarding p . In such cases of extended cognition, the most salient causal factor explaining true belief is not so much a cognitive trait of S as an external feature cognitively scaffolding her. Such scaffolding isn't creditable to S , hence strong credit theories of knowledge must be wrong.

Acknowledgements Research by Krist Vaesen was supported by the Netherlands Organisation for Scientific Research (NWO). He thanks Philip Nickel, Wybo Houkes, Martin Peterson, Anthonie Meijers, Duncan Pritchard, Andy Clark, Hans Radder, Joel Anderson, Lieven Decock, Olle Blomberg, Marieke van Holland, Auke Pols, Joel Katzav and Andreas Spahn for useful discussions on previous drafts of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14(1), 43–64.
- Adler, J. (2002). *Beliefs own ethics*. Cambridge, MA: MIT Press.
- Audi, R. (2001). Doxastic voluntarism and the ethics of belief. In M. Steup (Ed.), *Knowledge, truth, and duty* (pp. 93–113). Oxford: Oxford University Press.
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford: Oxford University Press.
- Clark, A. (2007). Mementos revenge: The extended mind, extended. In R. Menary (Ed.), *The extended mind*. Aldershot: Ashgate.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 1, 7–19.
- Greco, J. (2003). Knowledge as credit for true belief. In M. DePaul & L. Zagzebski (Eds.), *Intellectual virtue: Perspectives from ethics and epistemology* (pp. 111–134). Oxford: Oxford University Press.
- Greco, J. (2007). The nature of ability and the purpose of knowledge. *Philosophical Issues*, 17(1), 57–69.
- Greco, J. (2008). Whats wrong with contextualism? *Philosophical Quarterly*, 58(232), 416–436.
- Hieronymi, P. (2006). Controlling attitudes. *Pacific Philosophical Quarterly*, 87(1), 45–74.
- Kvanvig, J. (2009). Responses to critics. In D. Pritchard, A. Haddock, & A. Millar, *The value of knowledge*. Oxford: Oxford University Press.
- Lackey, J. (2007). Why we dont deserve credit for everything we know. *Synthese*, 158(3), 345–361.
- Lackey, J. (2009). Knowledge and credit. *Philosophical Studies*, 142(1), 27–42.
- Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology*, 19(3), 329–344.
- Owens, D. (2000). *Reason without freedom: The problem of epistemic normativity*. London: Routledge.
- Pritchard, D. (2005a). *Epistemic luck*. Oxford: Oxford University Press.
- Pritchard, D. (2005b). Virtue epistemology and the acquisition of knowledge. *Philosophical Explorations*, 8(3), 229–243.
- Pritchard, D. (forthcoming). Cognitive ability and the extended cognition thesis. *Synthese*.
- Raz, J. (1999). *Engaging reason: On the theory of value and action*. Oxford: Oxford University Press.
- Riggs, W. D. (1998). What are the chances of being justified?. *The Monist*, 81(3), 452–473.
- Riggs, W. D. (2007). Why epistemologists are so down on their luck. *Synthese*, 158, 329–344.
- Riggs, W. D. (2009). Two problems of easy credit. *Synthese*, 169, 201–216.
- Rowlands, M. (2009). Extended cognition and the mark of the cognitive. *Philosophical Psychology*, 22(1), 1–19.
- Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, CI(8), 389–428.
- Sosa, E. (2006). Knowledge: Instrumental and testimonial. In J. Lackey & E. Sosa (Eds.), *The epistemology of testimony* (pp. 116–126). Oxford: Oxford University Press.

- Sosa, E. (2007). *A virtue epistemology: Apt belief and reflective knowledge* (Vol. 1). Oxford: Oxford University Press.
- Vicente, K. J. (2003). *The human factor: Revolutionizing the way people live with technology*. Toronto: A.A. Knopf Canada.