JFGI: FROM DISTRIBUTED COGNITION TO DISTRIBUTED RELIABILISM[*]

Kourken Michaelian
Bilkent University

While, prima facie, virtue/credit approaches in epistemology would appear to be in tension with distributed/extended approaches in cognitive science, Pritchard (2010) has recently argued that the tension here is only apparent, at least given a weak version of distributed cognition, which claims merely that external resources often make critical contributions to the formation of true belief, and a weak virtue theory, which claims merely that, whenever a subject achieves knowledge, his cognitive agency makes a significant contribution to the formation of a true belief. But the significance of the role played by the subject's cognitive agency in distributed cognitive systems is in fact highly variable: at one extreme, formation of a true belief seems clearly to be significantly creditable to the subject's agency; at the other extreme, however, the subject's agency plays such a peripheral role that it is at best unclear whether it should receive significant credit for formation of a true belief. The compatibility of distributed cognition and virtue epistemology thus turns on what it takes for a contribution to the formation of true belief to count as significant. This article argues that the inevitable vagueness of this notion suggests retreating from virtue epistemology to a form of process reliabilism and explores the prospects for a distributed reliabilist epistemology designed to fit smoothly with distributed cognition. In effect, distributed reliabilism radicalizes Goldberg's recent extended reliabilist view (Goldberg 2010) by allowing the process the reliability of which determines the epistemic status of a subject's belief to extend to include not

---

only processing performed by other subjects but also processing performed by non-human technological resources.

"JFGI" is an acronym, occasionally used in online discussions, for "just fucking google it". The phrase is mostly used in response to questions the answers to which are easily obtainable via a simple web search (indeed, the questioner may simply be directed to http://justfuckinggoogleit.com/), the suggestion apparently being that, since it is just as good, from the questioner's point of view, to retrieve easily googleable information from a website as it is to receive it from other users, he ought to do so, thus sparing other users the expenditure of unnecessary effort. The phrase thus hints at a view on which reliance on other subjects and reliance on non-human resources are on a par, epistemically speaking, and it is towards such a view that I want to move in this article.[1]

I ultimately suggest a version of process reliabilism meant as an explicit statement of the view. Traditionally, reliabilists have viewed both the provision of testimony by human interlocutors and the provision of information by non-human technological resources as being mere background conditions on belief-formation, capable of influencing the reliability of a subject's belief-forming processes but not of constituting parts of those processes. Goldberg (2010) has recently moved away from the traditional view, arguing that testimony, in particular, may count as part of the cognitive process responsible for the production of a belief—that, in the case of testimonial belief-formation, the epistemically relevant process (the process the reliability of which determines the epistemic status of the target belief) extends beyond the believing subject to include processing performed by the communicator who provides the relevant testimony. However, he explicitly rejects the possibility of making an analogous move with respect to the contribution made by non-human resources, denying that, in cases in which the subject depends on information provided by non-human resources in a manner apparently analogous to that in which he depends on information provided by other humans in testimony cases, the epistemically relevant process extends beyond the believing subject to include processing performed by the relevant external resources (Goldberg 2012). Drawing on the distributed/extended cognition approach in cognitive science, my goal here is to motivate this more extreme departure from the traditional reliabilist view.

1. Distributed Cognition and Virtue Epistemology

The route that I take to this goal will be somewhat indirect. In recent years, many epistemologists have abandoned process reliabilism for one

or another form of virtue reliabilism. (While the forms of virtue reliabilism in question include production by a reliable process as a condition on knowledge, they go beyond simple process reliabilism by requiring that the reliability be attributable to the agent in some sense.) The arguments in favour of making this move are reasonably persuasive, and I therefore begin with virtue epistemology, looking at the current debate over the compatibility of distributed cognition and virtue epistemology. While, prima facie, virtue/credit approaches in epistemology would appear to be in tension with distributed/extended approaches in cognitive science, Pritchard (2010) has recently argued that the tension here is only apparent, at least given a weak version of distributed cognition, which claims merely that external resources often make critical contributions to the formation of true belief, and a weak virtue theory, which claims merely that, whenever a subject achieves knowledge, his cognitive agency makes a significant contribution to the formation of a true belief. But the significance of the role played by the subject's cognitive agency in distributed cognitive systems is in fact highly variable: at one extreme, formation of a true belief seems clearly to be significantly creditable to the subject's agency; at the other extreme, however, the subject's agency plays such a peripheral role that it is at best unclear whether it should receive significant credit for formation of a true belief. The compatibility of distributed cognition and virtue epistemology thus turns on what it takes for a contribution to the formation of true belief to count as significant. This article argues that the inevitable vagueness of this notion suggests retreating from virtue epistemology to a form of process reliabilism and explores the prospects for a distributed reliabilist epistemology designed to fit smoothly with distributed cognition. In effect, distributed reliabilism radicalizes Goldberg's recent extended reliabilist view (2010) by allowing the process the reliability of which determines the epistemic status of a subject's belief to extend to include not only processing performed by other subjects but also processing performed by non-human technological resources.

1.1. Context

Before proceeding, three brief clarificatory remarks: First, *distributed* cognition is often used to refer to the tradition of theoretical and empirical work in cognitive science, associated with Hutchins (Hutchins 1995, 1996; Hollan et al. 2000), which focusses on cognitive processing in sociotechnical systems including multiple human and technological components, while *extended* cognition is often reserved for the current in philosophy of mind, associated with Clark (Clark and Chalmers 1998; Clark 1997 2008), which tends to focus primarily on cognition in systems centred on a single human augmented by external technological resources. While there are certainly differences of emphasis between Hutchins and Clark (Blomberg 2009), these do

appear for the most part to be merely differences of emphasis, at least if we set aside strong metaphysical claims sometimes made by extended cognition theorists about external resources literally constituting parts of agents' minds (Michaelian and Sutton 2013), and I will therefore refer interchangeably to distributed and extended cognition (defaulting to "distributed").

Second, as my focus here is on the epistemic status of the beliefs of individual subjects, I focus primarily on systems centred on single subjects. It should be possible to extend my argument to beliefs formed by subjects embedded in large-scale distributed socio-technical systems, but I will not argue that knowledge can be attributed to such systems themselves.

Finally, the forms that cognitive extension can take are of course many and varied—Sutton cites, for example, external cultural tools, artifacts, and symbol systems; natural environmental resources; interpersonal and social distribution or scaffolding; embodied capacities and skills; and even internalized cognitive artifacts, e.g., mnemonics (Sutton 2006). With most of the literature on the compatibility of distributed cognition and virtue epistemology so far, I focus, relatively narrowly, on interactions between human subjects and either external technological resources or other human subjects. As far as the epistemology is concerned, there may be important differences between this group of cases and some of the remaining cases; in particular, I concede that, in cases where the relevant external resource is not engaged in information processing (roughly, in cases where it is not representational, in Heersmink's sense (Heersmink 2013))—for example, in cases involving natural environmental resources—there may not be the same pressure to extend the belief-forming process beyond the subject.

1.1.1. *Weak d-cog*

Setting aside differences between the traditions in philosophy and cognitive science, distributed/extended cognition can refer to a whole family of mutually incompatible views, ranging from strong views (such as those of Hutchins and Clark), on which the idea of a distributed or extended cognitive system is taken literally, to much weaker views (embedded cognition (Rupert 2009, 2013), scaffolded cognition (Sterelny 2010), situated cognition (Robbins and Aydede 2008; Roth and Jornet 2013), and so on), on which external resources make critical contributions to human cognition without thereby literally coming to count as parts of distributed/extended systems (see Smart 2010 for an overview). Most of the literature on the compatibility of distributed cognition and virtue epistemology so far focusses on a *very* weak view, which makes only the (by now) basically uncontroversial claim that "human cognition is strongly dependent on external resources (whether or not we call them cognitive)", as Vaesen puts it (Vaesen 2011, 521). Call this *weak d-cog*.

Since weak d-cog is an extremely moderate view, showing that virtue epistemology is incompatible with it would, from the point of view of opponents of the virtue approach, be a more impressive result than demonstrating incompatibility with a stronger view. By the same token, of course, showing that the virtue approach is *compatible* with a stronger distributed view would, from the point of view of proponents of the virtue approach, be a more impressive result than demonstrating compatibility with the weak view. However, the key point is that all parties to the debate will (or should) accept weak d-cog, disagreeing over how much, if at all, to strengthen it—weak d-cog is so weak that virtually all *opponents* of more robust distributed approaches should accept it. The stakes of the debate are thus extremely high. If the virtue approach turns out to be incompatible with weak d-cog, it is incompatible with a claim that essentially all researchers in both epistemology and cognitive science would accept. Incompatibility with weak d-cog would thus count as very strong evidence against virtue epistemology, providing good reason to abandon virtue epistemology.[2]

1.1.2. $COGA_{WEAK}$

According to virtue/credit epistemology in general, knowledge can be (partly) analyzed in terms of the degree to which the formation of the relevant true belief by the subject is due to his reliable cognitive character, i.e., the extent to which formation of a true belief is attributable to his cognitive ability or agency (agency and character are treated more or less interchangeably in this framework). Like distributed/extended cognition, virtue/credit epistemology refers to a broad family of views. On strong views (e.g., Sosa 2007; Zagzebski 1996; Greco 2007), knowledge can be wholly analyzed in terms of cognitive ability (that is, knowledge just is the product of cognitive ability). On weak views (e.g., Pritchard 2008 2012), in contrast, cognitive ability provides a necessary but not a sufficient condition on knowledge.

Most of the literature on the compatibility of d-cog and virtue/credit epistemology so far focusses on relatively weak views. Pritchard (2010), in particular, discusses the following view.

COGA$_{WEAK}$ If S knows that P, then S's true belief that P is the result of a reliable belief-forming process which is appropriately integrated within S's cognitive character such that his cognitive success is to a significant degree creditable to his cognitive agency.

COGA$_{WEAK}$ contrasts with COGA$_{STRONG}$, which differs from it in two respects: first, the conditional is replaced with a biconditional; second, the

subject's cognitive success (formation of a true belief) is said to be *primarily*, rather than significantly, creditable to the subject's cognitive agency.

Why prefer $COGA_{WEAK}$ to $COGA_{STRONG?}$ Pritchard offers two main reasons. First, $COGA_{STRONG}$ cannot handle certain types of epistemic luck. In the well-known "barn country" case (Goldman 1976), for example, the subject's cognitive success does seem to be attributable primarily to his cognitive agency, but nevertheless he lacks knowledge, due to the influence of epistemic luck (the luck here is said to be environmental, rather than intervening, as it is, e.g., in Chisholm's "sheep in the field" case (Chisholm 1977)). This suggests that an explicit anti-luck condition is required, i.e., that attributability of cognitive success to the subject's agency can be at most a necessary condition on knowledge. Second, even if we therefore break the biconditional, the view still needs to be weakened further, for there seem to be cases in which the subject knows without his cognitive success being *primarily* attributable to his cognitive agency. Pritchard offers a case (originally due to Lackey 2007) in which a subject needs directions in an unfamiliar city, asks the first person that he meets, and accepts the directions more or less uncritically. The idea is that the subject knows, but that his formation of a true belief, while due *to a significant extent* to his cognitive agency (because he chooses to ask someone who looks like he lives in the city, rather than another tourist, etc.), is not due *primarily* to his cognitive agency.[3]

1.2. The current state of play

Setting aside stronger versions of virtue epistemology, is weak d-cog compatible with a relatively weak view along the lines of $COGA_{WEAK}$? Pritchard argues that it is, claiming that, even in simple cases such as Clark and Chalmers' well-known Otto case (Clark and Chalmers 1998), the subject's cognitive agency makes a significant contribution to the formation of a true belief. According to Pritchard, Otto's acquisition of his notebook already "represents a great deal of epistemic virtue on his part":

> A lesser cognitive agent ... would have acquiesced in the loss of his (non-extended) memory and so accepted the epistemic consequences. Moreover, notice that the way in which Otto employs his notebook also reflects his epistemic virtue. An agent less concerned with epistemic goods would not, for example, go to the lengths that Otto goes to in order to ensure that this information resource is readily available to him but really would just use this notebook as a mere incidental aid to his cognition. (Pritchard 2010, 233)

In support of this claim, we might note, further, that Clark and Chalmers' description of Otto as automatically endorsing information retrieved from his notebook is somewhat misleading—in a more realistic version of the

case, Otto, as a cognitively virtuous subject, does not automatically endorse information retrieved from the notebook but rather is sensitive to signs that retrieved information has become outdated, that the notebook has been tampered with, etc., just as normal subjects monitor their own biological memories for cues to inaccuracy (Michaelian 2012a).

The key claim of the strategy, then, is that, due to Otto's active role in setting up and using his notebook, the notebook counts as being integrated into his cognitive character in a way that allows his cognitive success, when he forms a true belief by relying on the notebook, to be significantly creditable to his cognitive agency. If this is right, then, when Otto forms a true belief by relying on the notebook, the belief may, given COGA$_{\text{WEAK}}$, amount to knowledge.[4] In general, as long as, in every case of knowledgeable belief in which the subject relies, in the manner described by weak d-cog, on an external resource in forming the relevant true belief (including, for example, Lackey's tourist case), the subject's own cognitive agency plays a significant role in achieving the formation of a true belief, COGA$_{\text{WEAK}}$ will turn out to be compatible with weak d-cog.

The question at this stage of the game, then, is: is it the case that, whenever S comes to know that P by relying on an external resource in the manner described by weak d-cog, S's cognitive agency nevertheless plays a significant role in the formation of S's belief that P?

2. The Role of Agency

In order to work out an answer to this question, I suggest that we need to look at a broader range of cases than have been discussed in the literature so far, and to do so more systematically.

2.1. Agency in distributed systems

Pritchard mentions two ways in which Otto's agency contributes to his success: first, Otto is responsible for acquiring and setting up the notebook in the first place; second, Otto is responsible for maintaining his access to the notebook, updating the information that it contains, and so on. I noted a third way in which Otto's agency might play a role: he is responsible for deciding whether or not to accept information retrieved from the notebook. (I emphasize that we should not overintellectualize: the relevant decisions may be—and in real cases normally are—made automatically and unconsciously.) These can be viewed as particular instances of three general ways in which a subject's agency can contribute to his cognitive success when he forms a true belief by relying on external resources in the manner described by weak d-cog: cognitive agency can contribute to the *selection* of the relevant external

resources, the *assembly* of the distributed system, and the *endorsement* of information produced by the system.

2.1.1. *Selection and endorsement*

In a previous paper (Michaelian 2012b), I argued that agents who rely on imperfect information sources (i.e., all real agents) face what I referred to as the *endorsement problem*, the problem of forming beliefs in a reliable manner despite relying on imperfectly reliable information sources. My claim was that a capacity for metacognitive monitoring and control plays a crucial role in explaining how agents solve this problem. For example, due in part to the reconstructive character of episodic memory retrieval (Michaelian 2011), in which information from various sources is combined and transformed to produce the representation eventually output by "retrieval", an agent faces a real risk of endorsing inaccurate representations of his past experiences and hence of forming false memory beliefs. I argued that, in the case of episodic memory retrieval, agents solve the endorsement problem through the form of metacognitive monitoring described by the source monitoring framework (Johnson et al. 1993; Mitchell and Johnson 2009), essentially by relying on key content-based characteristics of retrieved information to infer source. I also briefly discussed what I referred to as the *selection problem*, the problem of selecting one's resources so that they provide one with answers to whatever questions are currently driving inquiry, taking into account the availability of resources, their reliability, the costs involved in their use, and so on. For example, in cases where an agent has previously determined the answer to a question, he might choose between retrieving the answer from memory and working it out anew (see section 2.2.1 below).

These initial formulations applied only to with *internal* versions of the endorsement and selection problems. As Arango-Muñoz has recently pointed out, however, selection and endorsement look somewhat more complicated when we take into account the role of external resources highlighted by weak d-cog. First, the selection problem now becomes what he refers to as the *extended* selection problem: "[s]ince normal subjects routinely use external as well as internal resources, each time they are confronted with a cognitive problem, they have to choose whether to solve it internally or externally" (Arango-Muñoz 2013, 139). With our increasing reliance on various online resources, the extended selection problem becomes increasingly prominent; when I need to retrieve information, do I rely on my own internal memory or do I JFGI? For example, Sparrow et al. (2011) (drawing in part on the transactive memory framework (Ren and Argote 2011; Theiner 2013)) have recently found that, when subjects expect to have access to information in the future, memory for that information decreases, while memory for where to find the information increases, an apparently efficient strategy for solving

this particular instance of the selection problem. While Arango-Muñoz's formulation of the extended selection problem is a step in the right direction, it should be emphasized that, in practice, solving the extended selection problem is often not a matter of choosing between mutually exclusive internal and external resources; especially in more complex tasks, the subject often selects a combination of internal and external resources. For example, one may rely on internal memory for some aspects of the needed information while relying on multiple external resources for additional aspects of the information.

Second, taking into account the role of external resources in cognition, the endorsement problem now has two distinct aspects. *Endorsement of internally retrieved information*: when the subject internally generates a representation, he must decide whether or not to endorse it. This corresponds to the formulation of the endorsement problem given in Michaelian 2012b. *Endorsement of externally retrieved information*: when the subject acquires information from an external resource, he must decide whether or not to endorse it. For example, after retrieving information from a website, the subject needs to decide whether to go ahead and endorse it (Chiu et al. 2013). Again, in practice, endorsement will often be more complicated than this formulation suggests. The representation produced by an extended process may incorporate information produced by a combination of internal and external resources, and the subject's decision whether or not to endorse it will presumably be sensitive to his individual levels of confidence in the various resources involved, as well as to his confidence in the particular combination of resources used, relative to the task at hand.

2.1.2. *Assembly*

The extended selection problem and the external version of the endorsement problem correspond to two of the ways in which Otto's agency makes a credit-worthy contribution to his formation of a true belief: he is responsible for selecting it (rather than continuing to rely on his declining biological memory), and he is responsible for choosing to endorse information retrieved from the notebook on particular occasions. The final way in which Otto's agency makes a contribution brings out that subjects who rely on external resources in the manner described by weak d-cog face a third problem, what we might refer to as the *assembly problem*: having selected an appropriate resource (or complex of resources), the agent must construct the distributed system, coordinate its activity, and monitor its functioning.[5]

Kirchhoff and Newsome usefully distinguish between two aspects of assembly. *Construction* refers to "the actual putting together of parts to form a distributed cognitive system" (2012, 166). The external components of distributed systems are not always ready-made or readily available. The agent

may need to act to ensure that he has access to the relevant resource; in many cases, he may need to modify an existing resource so that it suits his informational needs or even to create it more or less from scratch. *Coordination* refers to "the coordination and continuous maintenance of the system's parts" (Kirchhoff and Newsome 2012, 166). The agent may need to combine information from various resources, ensure that resources deliver their information in the right order, and so on. Additionally, the agent may need to monitor the assembled system (and its individual components) to ensure that they are functioning as desired. A given component may, for example, fail, or work too slowly, or turn out to be unable to provide the information it was expected to provide.

2.2. The variable contribution of agency

Considering a range of cases of interaction both with other human agents and with technological resources, it will emerge that distributed cognition is a mixed bag, as far as the contribution of agency is concerned. In some cases, it seems clear that agency makes the sort of significant contribution required by $\text{COGA}_{\text{WEAK}}$, whether in selection, assembly, or endorsement. In other cases, however—in particular, in cases involving straightforward reliance on testimony provided by unknown informants or analogous reliance on instruments—it is at best unclear whether the subject's agency makes a significant contribution at any stage of the belief-forming process. The compatibility of virtue epistemology and weak d-cog will thus turn out to depend on what it takes for a contribution to the achievement of cognitive success to count as significant.

2.2.1. Agency in selection

In a relatively simple instance of the extended selection problem, the agent must choose between relying on an internal resource and relying on an external resource to accomplish a given cognitive task. In a realistic version of Clark and Chalmers' case, Otto, for example, retains some use of his biological memory, and so faces a choice between relying on his biological memory to retrieve the answer to a given question and relying on his notebook.[6] Even in such a relatively simple case, the agent must assess trade-offs among speed, accuracy, and the various costs involved in using internal and external memory: as Arango-Muñoz points out (drawing on Kalnikaité and Whittaker's investigation of factors determining the choice between internal and external memory (2007)), "[a]n internal strategy is normally quicker but has cognitive costs, since it requires working memory and attention; moreover, it is less accurate: An external strategy, on the other

hand, is more accurate, but has sensory-motor costs and is normally less efficient; i.e. it takes longer" (Arango-Muñoz, 2013 139). Moreover, there is evidence that subjects successfully negotiate such trade-offs. For example, when required to choose between retrieving the answer to an arithmetic problem from memory (the internal resource) and calculating it using pen and scrap paper (the external resource), subjects apparently rely on their feeling of knowing (FOK) to predict whether they can retrieve the answer or need to rely on the external resource to calculate it; and the FOK reliably predicts ability to retrieve the answer from memory (Reder and Ritter 1992; Paynter et al. 2009).

In a more complex instance of the selection problem, the subject chooses not simply between relying on an internal resource and relying on an external resource but rather among different assemblages of multiple internal and external resources. In such cases, the trade-offs that must be negotiated are more numerous and varied, and the contribution of the subject's agency will be correspondingly greater.[7] For example, when driving an unfamiliar route, I may rely on my own sense of direction (grounded in an appreciation of various environmental cues), information provided by a GPS unit, hard-copy maps, information provided by my passengers (who themselves may have access to various external resources), and information provided by strangers I/we ask for directions (e.g., see Forlizzi et al. 2010 on navigation as a collaborative process).

At the other extreme, however, are cases in which little effort or intelligence goes into selecting the relevant external resources, and in which the role played by the subject's agency appears to be correspondingly minimal. Consider again Lackey's tourist case, as redeployed by Pritchard. Pritchard suggests that it is clear that the tourist who, upon arriving in an unfamiliar city, simply asks the first person she sees for directions and forms her belief on the basis of the information he provides has knowledge, as long as we assume, inter alia, that she "is suitably responsive to epistemically relevant factors—it is not as if, for example, she would ask someone who would clearly not be a good informant, and it is not as if she would believe whatever she was told, even if it was obviously false" (2010, 141); the same assumption suggests that her cognitive agency makes a contribution to her formation of a true belief that will count as significant in the sense of COGA_WEAK. This sort of optimism about the capacity of subjects to select reliable informants and filter out unreliable information is shared by Sperber and colleagues (Sperber et al. 2010), who argue, on the basis of evidence including work on children's preference for trustworthy informants (Mascaro and Sperber 2009; Clément 2010), that we are reasonably good at selecting reliable informants (I return to the claim about our sensitivity to unreliable information, also endorsed by Sperber et al., in section 2.2.2). As I have argued elsewhere, however, this optimism may be unfounded: that we have a preference for reliable informants does not mean that we are in fact good at selecting reliable

informants (Michaelian 2013b; but see Sperber 2013 for a response). More-over, it seems that the tourist may gain knowledge from her informant's testimony even if she is not good at selecting reliable informants, at least assuming that she is in a sufficiently epistemically friendly environment, con-taining relatively few unreliable informants (Michaelian 2013a). (I want to know what time it is, but I don't have a watch. I ask the first person I pass on the street for the time. He tells me, and I believe him. If he were, for some reason, incompetent or dishonest with respect to the time, I would still have asked him for the time. But this need not prevent me from coming to know, on the basis of the information he provides, what time it is. Cf. Gerken 2013 on the contribution of the environment to testimonial justification.) If so, this provides us with a type of case in which, as far as selection is concerned, the subject appears to attain knowledge without his cognitive agency making a significant contribution.

2.2.2. *Agency in endorsement*

In a relatively simple instance of the endorsement problem, the subject must determine whether to endorse information retrieved either from his own biological memory (the internal version of the problem) or from an external memory store (the external version of the problem). It seems clear that the subject's agency makes a significant contribution to solving the internal ver-sion of the problem; if so, then, in parallel external cases, it should be clear that the subject's agency makes a significant contribution to solving the ex-ternal version of the problem. Elsewhere (Michaelian 2012b), I have argued that, as far as memory is concerned, subjects solve the endorsement problem by relying on the form of metacognitive monitoring described by Johnson et al.'s source monitoring framework (SMF) (Johnson et al. 1993; Mitchell and Johnson 2009) (metacognitive monitoring may play a similar role in other sources of knowledge, including perception (Loussouarn et al. 2011) and reasoning (Cox and Raja 2011)). According to the SMF, records are typically not tagged with source information; instead, subjects infer source using heuristics relying on average content-based differences (e.g., level of sensory, contextual, semantic, and affective detail) among records originat-ing in different sources. While source monitoring is normally unconscious and automatic (roughly, a type 1 process (Evans and Stanovich 2013)), it clearly counts as appropriately integrated into the subject's cognitive char-acter, thus allowing us to assign a significant degree of credit for formation of a true belief to the subject's cognitive character in cases involving the internal version of the endorsement problem. Arango-Muñoz (2013) has re-cently argued for an alternative view on the role of metacognition in solving the internal version of the endorsement problem, assigning the main role to metacognitive or epistemic feelings (de Sousa 2008; Arango-Muñoz 2014;

Dokic 2014; Proust 2013) based on subpersonal monitoring of cognitive processing, rather than to heuristic monitoring of the sort described by the SMF, in which epistemic feelings do not figure explicitly; positive feelings (e.g., of confidence or rightness) or negative feelings (e.g., of uncertainty or error) guide the subject's decision to accept or reject retrieved internally retrieved information. This account may ultimately be compatible with the source monitoring account, since, while epistemic feelings are available to consciousness, they are themselves produced by unconscious heuristic monitoring (Koriat 2000; Arango-Muñoz and Michaelian 2014). Whether or not the accounts are compatible, their implications are similar as far as cognitive agency are concerned. Moreover, either metacognitive feelings or the sort of non-affective monitoring described by the SMF may play an analogous role in enabling subjects to solve the external version of the endorsement problem. For example, the subject may rely on a feeling of truth (determined by cognitive fluency, i.e., subjective ease of processing (Oppenheimer 2008)) to shape his decision whether to endorse or reject externally retrieved information (Reber and Unkelbach 2010). Whatever the mechanism responsible for determining endorsement/rejection, there will clearly be a class of cases of external endorsement in which the subject's cognitive agency receives significant credit for formation of a true belief.

In more complex instances of the endorsement problem, in which the subject must determine whether to endorse (and with what level of confidence) a representation produced by combining information retrieved from multiple internal and external resources, the significance of the contribution of the subject's agency is equally unproblematic. Consider the case of conversational remembering (Sutton et al. 2010), in which groups of two or more individuals remember together. As Hirst and Echterhoff point out (2012), such distributed remembering has both benefits (e.g., collaborative facilitation, in which the group as a whole recalls more than individuals recall alone (Weldon 2001) and costs (e.g., collaborative inhibition, in which an individual remembering in a group recalls less than he would recall alone (Rajaram and Pereira-Pasarin 2010)). Though they do not necessarily view this as a cost (Fagin et al. 2013), collaborative remembering can also lead to social contagion of memory (Roediger et al. 2001), in which memories spread from one person to another by means of conversational interaction, in some cases resulting in memory that does not correspond to experience. The example of social contagion brings to the fore the need for subjects embedded in distributed memory systems to actively monitor the sources of retrieved information, in order to ensure the accuracy of their own subsequent memories, and hence emphasizes the significant contribution of cognitive agency.

However, returning to the case of testimony, in which a subject relatively passively receives information from another agent, as opposed to actively participating in the construction of a shared memory, it appears that the subject's agency may play an insignificant role in achieving formation of a

true belief. As noted above, Pritchard (2010), like Sperber et al. (2010) (and cf. Fricker, 1995), is optimistic about the ability of recipients of communicated information to filter out unreliable information; this optimism appears to be unwarranted. Even setting aside Gilbert's influential view, on which acceptance of communicated information is automatic (with possible rejection being a subsequent, effortful step) (Gilbert et al. 1990, 1993), there is abundant evidence (Michaelian 2010, 2012c) that we are simply not good at detecting dishonesty on the part of communicators; the typical finding, in fact, is that we are barely better than chance (Vrij 2008). The upshot is that, while there will be cases in which the subject's agency plays a role, in many cases, when the subject forms a true testimonial belief, the truth of the belief is due entirely to the truth of the received testimony—had the received information been false (because dishonest), he would have accepted it anyway; and it seems mistaken, in such cases, to assign a significant role to the subject's agency in explaining formation of a true belief, as far as endorsement is concerned. Of course, more active monitoring may be triggered if the speaker displays obvious signs of apparent dishonesty or incompetence, but in straightforward cases monitoring by the recipient does not play a significant role. We thus have an important class of cases—note that the class will include not only cases of reliance on testifiers but also cases of analogous reliance on non-human instruments—in which the subject's agency plays a significant role in explaining his success in forming a true belief neither in the selection phase nor in the endorsement phase of the belief-forming process. There remains the possibility that agency plays a significant role in the assembly phase, to which I turn next.

2.2.3. *Agency in assembly*

In some instances of the assembly problem, the subject's agency plays a role that intuitively counts as significant. Consider again Otto and his notebook. As Pritchard emphasizes, Otto plays an active role in constructing and maintaining the extended system. This is typical of the use of notebooks as external memory aids, including by patients with normal internal memories. The subject must organize the contents of the notebook if they are to be useful (constructing the distributed system). And the contents are then often used as prompts to aid the subject to retrieve information from internal memory (which requires the subject to coordinate the system), rather than simply acting as an external substitute for the internal memory store (see Yeo's historical study (2008)) and Kalnikaité and Whittaker (2008) for a contemporary approach). More generally, as notebooks and more novel forms of external memory (smartphones, etc.) are typically used to augment rather than simply replace biological memory (Clowes 2008, 2013), assembly of the relevant distributed systems normally involves an active contribution by the agent.

In more complex instances, the significance of the contribution of the subject's agency is all the more evident. Dählback, Kristiansson, and Stjernberg (2013), for example, explore mnemonic strategies used by elderly adults coping with cognitive decline. They emphasize that such strategies often make use of both artifacts (lists, material reminders) and other agents, and often involve backups and failsafes in order to produce resilient distributed memory systems. In many cases, such distributed systems are constructed largely (though usually not only) by the subject, and depend on the subject's active monitoring to maintain coordination.

But returning again to cases at the other end of the spectrum, such as Lackey's tourist case and analogous cases of straightforward reliance on available instruments, the significance of the contribution of the subject's agency to the assembly phase is at best unclear. Indeed, in many of the relevant cases, "assembly" is essentially a trivial process—the tourist simply asks the first person she sees for directions. Thus, as far as the role of agency in distributed cognitive systems is concerned, we are faced with a range of cases. At one extreme, we have a class of cases in which the subject plays an active role in selecting the assemblage of internal and external resources to be used, constructing and coordinating the overall distributed system, and actively monitoring the functioning of the system in order to determine whether to endorse the information output by the system; here, the contribution of the subject's agency presumably meets the standard of significance set by $COGA_{WEAK}$. At the other extreme, however, we have a class of cases in which the subject's agency appears to make only a negligeable contribution to selection, assembly, and endorsement. In principle, minimal contributions of the subject's agency at each of the selection, assembly, and endorsement phases might add up to a significant overall contribution, but it seems likely that this does not occur in the sort of cases of straightforward reliance on external resources that I have described: when I ask someone for directions, form a belief about the current time by looking at a clock, or JFGI, it is unclear whether the minimal role played by my own agency is sufficient to count as significant, in the relevant sense.

3. Retreating from Virtue Epistemology

In this final section, I suggest that this point about the variable contribution of agency gives us good reason to retreat from virtue epistemology to a form of process reliabilism designed to accommodate d-cog.

3.1. Attempts to save virtue epistemology

Before doing so, however, I want to consider two important strategies designd to save virtue epistemology.

3.1.1. *The contextualist strategy*

Kelp (2013) argues that Pritchard's and Vaesen's arguments against the compatibility of robust virtue epistemology (e.g., $COGA_{STRONG}$) and weak d-cog turn on the fact that the notion of primary creditability and related notions are context sensitive: depending on which features of the situation are salient in the conversational context, Otto*'s true belief (where Otto* is like Otto, except that his wife plays a crucial role in setting up his notebook for him), for example, may or may not seem to be primarily creditable to his cognitive agency. Kelp further argues that standard responses to worries about the context-sensitivity of attributions of primary creditability can thus be invoked to defeat arguments for the incompatibility of robust virtue epistemology and weak d-cog. One obvious possibility here is to adopt a form of contextualism about knowledge (Greco 2007).[8] A defender of virtue epistemology might attempt to respond to the worry, developed here, about the compatibility of weak d-cog and even $COGA_{WEAK}$ by adopting Kelp's strategy: just as attributions of primary creditability are context sensitive, attributions of significant creditability are context sensitive; hence, whether the creditability of the subject's agency in the relevant class of cases counts as significant will depend on which aspects of those cases are conversationally salient; if we adopt a form of contextualism, this need not be a problem.

Contextualists may be content with this strategy; the rest of us should not be. I have nothing novel to say against contextualism here, but I point out that those of us who are committed to a vision of epistemology as investigating natural phenomena (e.g., Kornblith 2002), with objective boundaries in the world, independent of what may or may not be salient in a given context, will reject the contextualist strategy for saving (weak) virtue epistemology.

3.1.2. *The distributed credit strategy*

Assuming that we reject the contextualist strategy, one possibility that suggests itself, in light of the lack of an objective standard for determining whether the contribution of a subject's agency to his cognitive success counts as significant and therefore creditworthy, is to refrain from attempting to partition credit, instead assigning credit to the distributed system as a whole. A view in this vicinity is suggested by Green (2012, 2014):

> CREDIT FOR US If *S* knows that *P*, then the abilities that contribute to the formation and sustenance of *S*'s belief that *P* deserve a high degree of credit for *S* knowing *P* whether those abilities are contributed solely by *S* or also by other agents.

It might appear that Green's view avoids the problem of vagueness that befalls COGA$_{\text{WEAK}}$, since, in order to apply the view, we do not need to be able to say whether the agency of the subject whose true belief is in question deserves a high degree of credit for his cognitive success—there is no need to partition credit among the various subjects who contribute to that success. But CREDIT FOR US restricts credit to the *human* (agential) components of distributed systems (in this respect, it is something like a virtue-theoretic analogue of Goldberg's extended reliabilism, discussed in section 3.2 below), and this feature of the view means that it falls prey to a version of the problem of vagueness, for it requires us to be able to say whether the human components of the distributed system (vs. the technological components) deserve a high degree of credit. The same problem would obviously also arise for a version of the view which referred to a *significant* degree of credit.

We might therefore move to a fully distributed credit view, along the following lines.

> CREDIT FOR US AND THE MACHINES If *S* knows that *P*, then the abilities that contribute to the formation and sustenance of *S*'s belief that *P* deserve credit for *S* knowing *P*, whether those abilities are contributed solely by *S* or also by other human agents *and* non-human resources.

But can we make sense of assigning credit for cognitive success not only to the human subjects who contribute to it but also to the non-human resources (e.g., instruments) on which they rely?

To the extent that it makes sense to assign credit only where there is agency, it would appear that views like CREDIT FOR US AND THE MA-CHINES are non-starters. It might, however, be possible to work out a notion of distributed agency. The core idea would be that we view distributed systems themselves as manifesting agency. Some components of distributed systems may also themselves manifest agency, while others (viz., the non-human components) do not. But in the context of knowledge attributions, we focus on distributed systems as wholes, viewing them as manifesting a form of agency or ability, which then receives credit for the successful formation of a true belief. Something like this distributed agency view is suggested by Kirchhoff (inspired in part by Sutton's remarks on the possibility of a "deterritorialized cognitive science" (Sutton 2010, 213)): "Unlike FP [the alternative "fixed properties" view], DP [his "dynamic properties" view] does not assume, when having to explain the integration/assembly of cognitive systems, that the individual organism is the most active element. DP implies that assembly of cognitive systems is the result of richly dynamical and distributed elements, where there is no collapse into individualism like in FP" (Kirchhoff 2012, 288). The result is a non-individualistic conception of cognitive agency. For example, Kirchhoff and Newsome, considering research on collaborative recall (Sutton et al. 2010), contend that, in cases of

collaborative recall, the distributed system composed by the collaborating subjects is "the appropriate target of epistemic credit" (Kirchhoff and Newsome 2012, 174). A fully distributed notion of agency, on which agency belongs to distributed systems as wholes (including both human and non-human components), would allow us to make sense of a view like CREDIT FOR US AND THE MACHINES and thus to retain a virtue approach without falling prey to the worry about vagueness that afflicts views like COGA$_{WEAK}$ and CREDIT FOR US.[9]

But the reasons offered by Kirchhoff for moving to a fully distributed conception of cognitive agency are far from decisive. Moreover, there are good reasons to be cautious about adopting such a conception. As Giere points out,

> The culture in scientifically advanced societies includes a concept of a human agent. According to this concept, agents are said to have minds as well as bodies. Agents are conscious of things in their environment and are self-conscious of themselves as actors in their environment. Agents have beliefs about themselves and their environments. Agents have memories of things past. Agents are capable of making plans and sometimes intentionally carrying them out. Agents are also responsible for their actions according to the standards of the culture and local communities. And they may justifiably claim to know some things and not other things. (Giere 2007, 316)

The sorts of non-human technological resources involved in distributed cognitive systems obviously do not satisfy these requirements. We can of course make sense of treating a system not all of the components of which are agents as being itself an agent (humans are agents, though their components are not themselves agents). When it comes to distributed systems, however, given that, to the extent that we need to refer to agency, we can appeal to the agency of their human components, parsimony dictates refraining from assigning agency to distributed systems as wholes (though we may still be able to motivate a view of agency as being distributed across the *human* components of distributed systems, as in CREDIT FOR US). Hence it is preferable to reject a fully distributed conception of agency, and with it views like CREDIT FOR US AND THE MACHINES, at which point we are left only with views like COGA$_{WEAK}$, which secure compatibility with d-cog at the cost of unacceptable vagueness.

3.2. Prospects for distributed reliabilism

Given that problems arise for virtue theories when we try to secure their compatibility with d-cog (though, if Kelp is right, these problems may ultimately be more general), a move that may be worth exploring is retreating

from virtue epistemology to a view closer to the sort of process reliabilism from which it descends. In the remainder of this section, I explore the prospects for this move, considering first Goldberg's view and then an alternative, distributed reliabilist view. Retreating from virtue/credit epistemology to process reliabilism involves certain costs, since we give up the advantages of virtue/credit theories that prompted the move in the first place; I consider these costs as well.

3.2.1.   *Extended belief-forming processes*

While we might, of course, attempt to combine d-cog with traditional Goldman-style process reliabilism, Goldberg has recently argued that belief-forming processes extend to include processing performed by other agents (though not processing performed by non-human external resources).

Goldberg's argument for the view that the process the reliability of which is relevant to the epistemic assessment of a belief may extend beyond the subject to include information processing performed by the testifier turns on the claim that the following "generic epistemic extended mind hypothesis" is true of testimonial belief (Goldberg 2012):

> GEEM For at least some cases in which a subject S believes that P, a proper epistemic assessment of S's belief requires an epistemic assessment of information processing that takes place in the subject's environment.

The argument for applying GEEM to testimony is essentially that it is implied by a (plausible) more general principle:

> when a subject S's belief that P is formed (or sustained) through a process π that takes as its input the output of a given stretch of cognitive processing π^*, then we should regard the belief-forming and -sustaining process relating to S's belief that P as including both π and π^*. (Goldberg 2012)

Memory provides one special case of this general principle: when assessing a memory belief, we consider not only the reliability of the process by which the belief was stored and retrieved but also the reliability of the process by which it was initially produced.[10] Memory is, in effect, a *temporally* extended process. Testimony, he argues (I return to his argument in section 3.2.2), then appears as another special case, with the extension now being *interpersonal* rather than temporal.

While it might seem that, if the process responsible for the formation of beliefs based on information received from human speakers should be characterized as extended, then so should the process responsible for the formation of beliefs based on information received from non-human

resources, Goldberg argues that GEEM cannot be applied in this way. His argument for this negative claim has two steps. First, he argues that "to rely in belief-formation on another speaker is to rely on an epistemic subject, that is, on a system which itself is susceptible to epistemic assessment in its own right, whereas "mere" instruments are mechanisms are not properly regarded as epistemic subjects in their own right, they are not susceptible to normative epistemic assessment" (2012, 182). Second, he argues that this (supposed) disanalogy between humans and instruments makes a difference to how we should delineate belief-forming processes. I deal with these two steps in turn.

Step 1: Are non-human components of distributed systems subject to normative epistemic assessment? Invoking a distinction originally drawn by Goldman (1979), Goldberg argues that we should classify the contributions of (non-human) external resources to cognitive processing as "brute-causal", as opposed to "cognitive-psychological". Now, we do indeed need some way of distinguishing between those causal antecedents of a belief that are relevant to its epistemic status and those that are not, since counting the entire causal history as relevant would lead to counterintuitive verdicts (as Goldberg points out) or even make epistemic evaluation impossible; Goldberg's proposal is that we should say that only processing performed by human subjects is relevant here, for human subjects, unlike other external resources, are subject to normative assessment, in the sense that they can display (or fail to display) epistemic responsibility and rationality, as opposed to mere reliability (Goldberg 2012, 188). For example, he suggests, what goes wrong in Bonjour's well-known case of Norman the clairvoyant (Bonjour, 1980) is that Norman lacks responsibility (despite having a form of reliable clairvoyance).

The sort of response to reliable clairvoyance and similar cases that Goldberg invokes here is not the only sort of response available—one might continue to defend a simple form of reliabilism, which denies that factors such as responsibility are relevant to the normative status of beliefs. For example, as Kornblith points out (2012), requiring, as Bonjour does, that the subject reflect on the epistemic status of his beliefs before he can be said to have knowledge sets the standards for knowledge extremely high; we may well want to resist this move—to say that (perhaps counterintuitively) the reliable clairvoyant's beliefs are not epistemically defective. But there is a more serious problem with Goldberg's claim that instruments are not subject to normative assessment; simply put, they are indeed subject to such assessment. The difference between instruments and agents is that normative assessment of the latter is sensitive to a broader range of factors than is normative assessment of the former, not that the latter, but not the former, are subject to normative assessment at all. In the case of instruments, normative assessment tends to be sensitive primarily to relatively thin properties, such as reliability. In the case of agents, normative assessment tends to be sensitive to (in addition to these thin properties) various thicker properties, such as rationality and responsibility. But that does not make assessment of instruments any less

normative, just as our assessment of agents does not become less normative when we take only their reliability into account.

This conclusion is sufficient to undermine Goldberg's argument against granting that GEEM applies to beliefs formed by reliance on non-human external resources. But even if we continue to insist that instruments are not subject to normative epistemic assessment, step 2 of the argument is equally problematic.

Step 2: Does extension depend on normative assessability? Goldberg's argument for the claim that only in cases where the external information source is itself subject to normative assessment should belief-forming processes be said to be extended is straightforward. Many processes carry information, and thus "we face a choice: we must either grant that Goldman's distinction between the brute-causal and the cognitive-psychological is irrelevant to doxastic justification after all; or else we must find some criterion to distinguish, from among all the information processing that is done in the world, which sorts are such that it is the "goodness" of those sorts that is relevant to epistemic assessments of doxastic justification" (2012, 188). For reasons given above, the former possibility is unacceptable; and Goldberg suggests that his normative assessability criterion is the best criterion available. However, he provides little argument for this. Moreover, the criterion faces two problems.

First, it is unacceptable on broad methodological grounds. Employing normative assessability as the criterion for the extension of belief-forming process is naturalistically unacceptable. From a naturalistic point of view, the question whether a given belief-forming process is extended or not—the question of how to delimit the process—should turn on features of the process itself, not on whether we are prepared to assess it in normative terms. (Of course, if we adopt a sufficiently deeply naturalistic account of epistemic norms, such as that developed by Proust (2013), then this worry about the naturalistic acceptability of the normative assessability criterion may go away. Adopting such an account, however, simply reinforces the point that instruments as well as agents are subject to normative assessment.)

Second, there are plausible, naturalistically-acceptable alternative criteria available. The need for a criterion arises, recall, because we need to distinguish between those external information-carrying factors that should be counted as part of the belief-forming processes and those external information-carrying factors that should be excluded, counting instead as part of the mere causal background of belief-formation. One obvious way of drawing this distinction is to say that external factors that merely carry information (e.g., a tree that conveys information about its age via the number of its rings) are part of the mere causal background, whereas external factors that actually engage in information processing (e.g., an instrument designed to calculate the age of a tree) can count as part of the belief-forming process. Another way is provided by Palermos, who argues that the sort of continuous mutual interaction that is described by dynamical systems theory provides a

criterion for extension of belief-forming processes (Palermos 2011, 2014). Or we might want to appeal here to the notion of cognitive integration. It is not my aim here to provide an argument for selecting one of these criteria; my point is simply that there are better criteria available. And these alternative criteria do not have the implication that Goldberg needs them to have in order to resist extending belief-forming processes into non-human resources, in addition to other human agents.

3.2.2. *Distributed belief-forming processes*

If the argument of section 3.2.1 is right, then, given that we should say that belief-forming processes extend to include processing performed by speakers in cases of testimony, we should also say that belief-forming processes extend to include processing performed by non-human resources where these play a role analogous to that played by speakers—roughly, transmitting the information that, if endorsed by the subject, constitutes the content of the resulting belief.[11] (In order to have a label to distinguish between Goldberg's "speakers-only" view and the alternative "speakers + instruments" view, I will refer to the former as *extended* reliabilism and to the latter as *distributed* reliabilism.) But what positive reason do we have for extending the process in the case of testimony, in the first place?

Goldberg's positive argument, as noted above, involves an appeal to a view of memory as a temporally extended process; the idea is that memory is a reliable *belief-dependent* process (in approximately Goldman's sense (1979)). One might attempt to resist extending the memory process by making the epistemic status of the memory belief depend only on the subject's current beliefs about whether he genuinely remembers or not, but, as Goldberg points out (2010, 70), this will not work: if the process by which the subject initially formed the relevant belief was unreliable, the current belief is unjustified, whatever the subject's beliefs about what he remembers. Thus, when a subject forms a belief by retrieving a belief from memory, the process the reliability of which determines the epistemic status of the belief extends (temporally) to include the earlier process (for example, perception) by which the belief was initially produced. Something similar seems to go for testimony: we should not say that only the subject's beliefs about the relevant testimony influence the epistemic status of his testimonial belief; if the testimony is unreliable, this implies that the resulting testimonial belief is unjustified, whatever the subject's beliefs about the testimony in question.

In order to count testimony as a belief-dependent process, Goldberg obviously needs to liberalize the concept of a belief-dependent process (it is not as if the communicator's own beliefs are the inputs to the process). He proposes the following intuitive liberalization:

> First, a belief-dependent process [on the liberalized understanding] must be a cognitive process that the reliability of whose outputs are a function of the reliability of its inputs. This, in turn, requires that the process should have inputs, and that these inputs be assessable in terms of their reliability. And, in order to be strict about what it is for an input to be assessable in terms of its reliability, we will insist that an input satisfies this condition only if it (the input) is itself the output of a cognitive process (process-type) whose reliability can be assessed in *its* turn. (Goldberg 2010, 72)

This liberalized concept of a belief-dependent process (a "quasi-belief dependent" process) allows us to count testimonial belief formation as a belief-dependent process, since the reliability of testimonial belief formation depends on the reliability of the relevant testimony.

Similarly, the concept allows us—as long as we do not interpret the requirement that the input be the output of a *cognitive* process in a question-begging way—to count belief-formation relying on information received from non-human resources in a manner analogous to reliance on testimonial information as a quasi-belief dependent process, despite the fact that non-human resources presumably do not themselves have beliefs in any strict sense. Consider a pair of simple cases:

Case 1: Subject S, who has been inside all day, asks agent A, who has just come in from outside, what the outside temperature is. A says that it is about 35°C outside. S endorses the information received from A and thereby forms a belief that it is about 35°C outside.

Case 2: S, who has been inside all day, looks through the window at a thermometer T attached to the building. T indicates that it is about 35°C outside. S endorses the information received from T and thereby forms a belief that it is about 35°C outside.

Of course, one might insist that, in case 2, only the subject's beliefs about the reliability of the thermometer influence the epistemic status of his temperature belief, whereas, in case 1, the reliability of the information-processing performed by A is relevant. But, intuitively, the epistemic status of that belief varies with the reliability of the thermometer, just as the epistemic status of the subject's belief in case 1 varies with the reliability of the testifier. Thus, absent strong reason not to, we should resist relegating the workings of the thermometer to the mere causal background against which the belief-forming process unfolds and should instead include the thermometer's contribution as part of the belief-forming process. I have argued above that Goldberg's attempt to give us such a reason is unconvincing. Thus I conclude (provisionally) that, given that we should move to extended reliabilism, we should go one step further and move all the way to distributed reliabilism.

Goldberg's argument for extended reliabilism, on which I rely here, has of course been subject to criticism. While I do not have space here to engage with this criticism in any detail, my argument here is essentially for the conditional claim that *if* we should move to extended reliabilism, then we should move to distributed reliabilism. Depending on how the debate unfolds, it may turn out that the initial move to extended reliabilism is insufficiently motivated, or cannot be motivated by the arguments provided by Goldberg. But I note here that I am more optimistic than some of the critics. For example, Gerken suggest that the notion of a quasi-belief dependent process may overliberalize the notion of belief-dependence, with the result that even perceptual beliefs may end up qualifying as (quasi-)belief dependent, since they depend on outputs of reliability-assessable perceptual processes (Gerken 2012) (cf. Malmgren 2011). However, in light of the cognitive penetrability of perception (Deroy 2013; Stokes 2013), we may already have reason to view the formation of perceptual beliefs as being (strictly) belief dependent, so I do not take the worry about overliberalization to be particularly pressing.

I emphasize that, by urging a move to distributed reliabilism, I do not mean to suggest that the internal structure of extended belief-forming processes should be disregarded. As Gerken's hypothesis of outsourced cognition (Gerken 2014), for example, emphasizes, the offloading of cognitive processing to external resources is itself a cognitive process, and the interface between internal and external resources will often be explanatorily important. I acknowledge this point, but I do not take it to count against distributed reliabilism, for, just as nothing prevents us from taking account of the complex structure of belief forming processes that are entirely internal, nothing prevents us from taking account of structure in the case of extended processes. For example, as noted above, epistemologists sometimes describe remembering as a simple belief-dependent process, taking a belief as input, storing it, then producing the same belief as output. In reality, even if we bracket social influences on memory (Sutton et al. 2010; Stone et al. 2012; Michaelian 2013a), remembering is an extraordinarily complex process, pulling in, integrating, and transforming information from a variety of sources at every stage from initial encoding, through consolidation of a stable representation, to the reconstructive retrieval process that generates the content of the potential memory belief and the metacognitive monitoring processes that determine whether or not the generated representation ends up being believed (and, if so, with what level of confidence) (Michaelian 2011, 2012b). The fact that we refer, for purposes of epistemic assessment, to a single overall memory process does not prevent us from taking this complexity into account. Analogously, the fact that we refer to a single overall belief producing process in many cases of testimony- or instrument-based belief in no way prevents us from taking account of the role

of the agent's cognition in offloading processing onto the relevant external resource.

3.3. Potential problems for distributed reliabilism

While distributed reliabilism fits better with distributed cognition than do the alternative views considered here, it is not without problems.

3.3.1. The costs of retreating from virtue epistemology

Retreating from virtue epistemology to a form of process reliabilism has costs. In particular, we must give up on the idea—what Pritchard (2005) refers to as the ability intuition—that knowledge presupposes more than mere reliability, that reliability must result from a cognitive achievement on the part of the agent.

Consider the case of "Temp". Temp forms his beliefs about the temperature in the room by consulting a thermometer. The thermometer appears to be normal, but it is actually malfunctioning, and randomly indicates different temperatures. However, there is a hidden agent who controls the temperature in the room, adjusting it so that it matches the temperature indicated by the thermometer at any given time. Temp's temperature beliefs are reliably formed, but intuitively he does not *know* the temperature in the room, and one natural way of explaining why Temp lacks knowledge is to point to the fact that the reliability of his belief-forming process has nothing to do with his cognitive ability.

Interestingly, the implications of distributed reliabilism might diverge from those of standard process reliabilism here. Pritchard's description of the case presupposes that the reliability of the thermometer itself is irrelevant to the epistemic status of Temp's beliefs. But if we allow the belief-forming process to include the thermometer's functioning, the process as a whole will be unreliable, yielding the intuitively correct verdict that there is something epistemically defective about Temp's beliefs about the temperature in the room. However, distributed reliabilism cannot account for our intuitions in all cases. Consider the case of "Alvin" (originally due to Plantinga (1993), adapted by Pritchard (2010)), who has a brain lesion that randomly but reliably causes him to form true beliefs about arithmetical sums. Assuming that Alvin has no understanding of the nature of the lesion, the relevant beliefs intuitively fail to qualify as knowledge, yet reliabilism will count him as having knowledge. There are reliabilist responses to such cases available (e.g., Becker 2013), but I do not pretend to be able to show here that one of these responses works. I simply note that distributed reliabilism is no worse off than is standard process reliabilism, and that the debate over whether

cases such as the brain lesion case actually require us to move to virtue epistemology is ongoing. While the ability intuition is powerful, it is not decisive, and hence the benefits of distributed reliabilism may still outweigh its costs.

3.3.2. *Vagueness and reliabilist knowledge*

My argument for moving from virtue epistemology to distributed reliabilism depends on the problem of vagueness encountered by virtue theories when weak d-cog is taken into account. Discussing this problem, Green grants that, while "[t]here must be some minimal threshold of ability that an agent contributes in order to know on a credit view", "establishing where that minimal threshold should be placed is not easy" (Green 2014). However, he points out that many epistemologies face similar problems—for example, it is well-known that reliabilists cannot identify a precise threshold of reliability that a process must meet in order for a true belief that it produces to count as knowledge. This point is directly relevant to my strategy here: if Green is right, reliabilism (including distributed reliabilism) is afflicted by the same sort of vagueness (though for a different reason) that leads me to reject virtue epistemology.

I grant, of course, that reliabilism cannot specify a non-vague threshold of reliability that a process must cross in order to generate knowledge. One obvious move to make at this point, as Green points out, is to go contextualist, letting context determine the degree of reliability required for knowledge. Given that I reject making the analogous move to save COGA$_{WEAK}$ (section, 3.1.1), obviously I reject making this move to save reliabilism. What I want to suggest, instead, is that, rather than giving us a reason for rejecting reliabilism, the impossibility of specifying a non-vague threshold of reliability gives us a reason for abandoning the attempt to theorize *knowledge*, as opposed to other epistemically relevant properties, such as level of justification. This suggestion is of course highly unorthodox, and I can do no more than sketch it here. Given the history of failure of attempts to theorize knowledge, however, it should at least be granted that the suggestion is non-crazy.

Given a commitment to a form of naturalism that views epistemology as being in the business of theorizing natural phenomena, with objective boundaries, what should reliabilists say in response to the apparent impossibility of specifying a principled boundary between reliability insufficient for knowledge and reliability sufficient for knowledge? One possibility is to insist that it will prove possible, in the long run, to specify such a boundary. But this view seems to be more or less hopeless. Another possibility is to grant that we will be unable do so, but to argue that, rather than suggesting that there is something wrong with reliabilism, the impossibility of specifying a principled boundary suggests that the attempt to do so is itself mistaken, for

there is no such boundary in the world. Though we might conventionally refer to "knowledge", knowledge is not a state in the world distinct from true belief. Subjects have beliefs, and those beliefs can be true or false—there is a difference between believing and failing to believe, between believing accurately and believing inaccurately. True beliefs can be more or less reliably produced (among other epistemically relevant properties (Alston 1993)), but it is not as if, once some threshold of reliability is crossed, they become states of a different sort—knowledge, as opposed to mere true belief. Hence, while we may wish, for practical purposes, to specify a desirable degree of reliability in a given context, such specifications should not be taken to be attempts to mark a real difference between different kinds of true belief (knowledgeable vs. non-knowledgeable). It remains possible to evaluate true beliefs as being epistemically better or worse, since they may be more or less reliably produced, more or less sensitive, etc. But the differences here will be differences of degree, not differences of kind.

Note that such a "nihilist" strategy is unavailable to the virtue epistemologist, since the point of moving to virtue epistemology in the first place is to provide a theory of *knowledge* that can cope with certain hard cases. By the same token, adopting the nihilist strategy has the added benefit that it eliminates most of the costs involved in retreating from virtue epistemology to reliabilism, since the (supposed) advantages of virtue epistemology over reliabilism mostly have to do with the difficulty that reliabilism has in discriminating knowledgeable true beliefs from non-knowledgeable true beliefs in these cases.

Summing up: Pritchard has argued that a relatively weak version of virtue epistemology is compatible with weak d-cog, but whether this is so depends on whether, in cases of distributed cognition of the sort described by weak d-cog, the subject's cognitive agency in general plays a significant role in bringing about formation of true belief, and the role played by the subject's agency in fact is highly variable, with the contribution of agency being quite minimal in some cases. Even a minimal contribution may be enough to count as significant, but the notion of significance at work here is vague. Hence it may be preferable to move from virtue epistemology to reliabilism, thus avoiding any reference to the significance of the role of agency, and Goldberg's arguments for extended reliabilism can be developed further to ground a form of distributed reliabilism which fits particularly well with d-cog. It may be, in the end, that there is no epistemically significant difference between what happens when we ask someone for directions and what happens when we JFGI. While distributed reliabilism will not be to everyone's taste—in particular, those who are wedded to the ability intuition will reject it, as will those who insist that the task of epistemology is to provide us with a theory of knowledge rather than a theory of justification—I take the theory to be sufficiently well-motivated to merit further investigation.

Notes

1. Of course, websites are typically (though not always) written by human authors; googling may thus look like a indirect way of accessing testimony, in which case the phrase does not suggest the sort of view that I have in mind. But search engines are not neutral conduits for information, and so we may take the phrase to suggest more than the idea that indirect human testimony is epistemically on a par with direct human testimony. When you google something, you rely on the search engine to select relevant websites, filter out unreliable results, and so on. Indeed, in some cases, the search engine itself may attempt to answer your query (as when one searches on a phrase such as "temperature in Ankara", "local time in Sofia", etc.). At any rate, the same point could be made using the phrase "just fucking Wolfram Alpha it" (though Google confirms that no one has ever actually used this phrase), in reference to the "computational knowledge engine" (http://www.wolframalpha.com/) that attempts to directly compute answers to questions on the basis of information available online, rather than simply providing the user with links to pages containing relevant information.

2. Or at least standard virtue epistemology. It might be possible to retain the broad virtue approach by moving to a distributed credit view; I return to this possibility below.

3. Vaesen (2011) similarly argues that d-cog is incompatible with strong virtue theories, using the case of a baggage inspector:

> 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.

As Vaesen sets up the case, the relevant counterfactual situation is one in which Sissi uses SYSTEM1 and therefore forms a false belief. So the most salient factor explaining her formation of a true belief is the external resource, SYSTEM2—her cognitive success is not primarily creditable to her agency. As far as I can tell, any differences between Lackey's testimony case and Vaesen's baggage inspector case do not affect my argument here.

4. This strategy requires saying something about the conditions under which a process counts as integrated in the right way into an agent's cognitive character. Pritchard mentions several ways: the agent knows that and why the process is reliable; the process is (roughly) innate; the agent has been successfully relying on the process for a sufficiently long span of time. Of these criteria for cognitive integration, the first is extremely strong and is unlikely to be satisfied by many processes, the second will not be satisfied by many of the processes at issue in cases of cognitive extension, and the third will not be satisfied by newly acquired processes. Thus at this point we lack a unified account of cognitive integration; but an intuitive understanding should be enough for present purposes.

5. The language of "assembly" is used by Clark, e.g., in his Principle of Ecological Assembly ("the canny cognizer tends to recruit, on the spot, whatever mix of problem-solving resources will yield an acceptable result with a minimum of effort" (Clark 2008, 13)). As Hutchins notes, Clark emphasizes *assemblies* in the sense of assembled resources, rather than assembly as the *process* of assembling resources (Hutchins 2011). My focus here is on assembly as a process.

6. Consider the case of the well-known patient H.M., who developed exceptionally thorough retrograde and anterograde amnesia after surgery on his medial temporal lobe (intended to cure his epilepsy); even H.M., who did rely on external memory stores to compensate for his severely impaired memory, retained some ability to acquire new procedural memories as well as some (limited) use of his biological semantic memory (Corkin 2013).

7. As noted in section 1.2, we lack a satisfactory criterion of the cognitive integration of extended processes. This is particularly evident when considering cases where the agent puts together a novel assemblage of resources. Such processes will likely fail to satisfy the three criteria suggested by Pritchard (known reliability, innateness, and long-term use), yet, intuitively, they count as cognitively integrated, given the active role of the subject in creating the assemblage.

8. Kelp also considers adopting an alternative form of robust virtue epistemology, due to Sosa (2007), according to which what matters is whether the subject's cognitive success "manifests" ability on his part. I will not consider this alternative form of robust virtue epistemology here, though I suspect that a similar problem arises for it (how much of a contribution need the subject make before his success "manifests" ability?).

9. An additional advantage of CREDIT FOR US AND THE MACHINES is that it is more likely to be compatible with strong d-cog than are more traditional virtue approaches. Strong d-cog views the subject's internal cognitive resources as components of an extended cognitive system in a strict sense, on a par with the relevant external components (i.e., with the components of the extended system that are external from the subject's point of view), and hence would seem to acknowledge no principled reason for singling out the subject's agency when assigning credit for cognitive success.

10. I set aside the fact that this characterization of remembering as a matter of storing and retrieving beliefs, though common in epistemology, is fundamentally mistaken (Michaelian 2012a); it is sufficient for Goldberg's argument that, if memory were to function this way, then it would provide a special case of the general principle.

11. This way of describing the view may suggest that it assumes the truth of a strong form of the hypothesis of extended cognition; as a strictly epistemological view, however, it is meant to be neutral between stronger and weaker versions of distributed/extended cognition.

References

Alston, W. P. Epistemic desiderata. *Philosophy and Phenomenological Research* 53(3): 527–551, 1993.

Arango-Muñoz, S. Scaffolded memory and metacognitive feelings. *Review of Philosophy and Psychology* 4(1): 135–152, 2013.

Arango-Muñoz, S. Metacognitive feelings, self-ascriptions and mental actions. *Philosophical inquiries* 2(1): 145–162, 2014.

Arango-Muñoz, S. and K. Michaelian. Epistemic feelings, epistemic emotions: Review and introduction to the focus section. *Philosophical inquiries* 2(1): 97–122, 2014.

Becker, K. Why reliabilism does not permit easy knowledge. *Synthese* 190(17): 3751–3775, 2013.

Blomberg, O. Do socio-technical systems cognise? In: *Proceedings of the 2nd symposium on computing and philosophy*. Society for the Study of Artificial Intelligence and Simulation of Behaviour, 2009.

Bonjour, L. Externalist theories of empirical knowledge. *Midwest studies in philosophy* 5(1): 53–74, 1980.

Chisholm, R. *Theory of knowledge*. Prentice-Hall, 1977.

Chiu, Y.-L., J.-C. Liang, and C.-C. Tsai. Internet-specific epistemic beliefs and self-regulated learning in online academic information searching. *Metacognition and learning* 8(3): 235–260, 2013.

Clark, A. *Being there: Putting brain, body, and world together again*. MIT, 1997.

Clark, A. *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford, 2008.

Clark, A. and D. Chalmers. The extended mind. *Analysis* 58(1): 7–19, 1998.

Clément, F. To trust or not to trust? children's social epistemology. *Review of philosophy and psychology* 1(4): 531–549, 2010.

Clowes, R. W. Hybrid memory, cognitive technology, and the self. In: Y. Erdin and M. Bishop, eds., *Proceedings of AISB/IACAP World Congress*. 2008.

Clowes, R. W. The cognitive integration of e-memory. *Review of philosophy and psychology* 4(1): 107–133, 2013.

Corkin, S. *Permanent present tense: The unforgettable life of the amnesic patient, H.M.* Basic Books, 2013.

Cox, M. T. and A. Raja, eds. *Metareasoning: Thinking about thinking*. MIT, 2011.

Dählback, N., M. Kristiansson, and F. Stjernberg. Distributed remembering through active structuring of activities and environments. *Review of philosophy and psychology* 4(1): 153–165, 2013.

de Sousa, R. Epistemic feelings. In: U. Doğuoğlu and D. Kuenzle, eds., *Epistemology and emotions*. Ashgate, 2008.

Deroy, O. Object-sensitivity versus cognitive penetrability of perception. *Philosophical studies* 162(1): 87–107, 2013.

Dokic, J. Feelings of (un)certainty and margins for error. *Philosophical inquiries* 2(1): 123–144, 2014.

Evans, J. S. B. T. and K. E. Stanovich. Dual-process theories of higher cognition. *Perspectives on psychological science* 8(3): 223–241, 2013.

Fagin, M. M., J. K. Yamashiro, and W. C. Hirst. The adaptive function of distributed remembering: Contributions to the formation of collective memory. *Review of philosophy and psychology* 4(1): 91–106, 2013.

Forlizzi, J., W. C. Barley, and T. Seder. Where should I turn: Moving from individual to collaborative navigation strategies to inform the interaction design of future navigation systems. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, CHI '10, 1261–1270. New York, NY, USA: ACM, 2010.

Fricker, E. Telling and trusting: Reductionism and anti-reductionism in the epistemology of testimony. *Mind* 104(414): 393–411, 1995.

Gerken, M. Critical study of Goldberg's Relying on others. *Episteme* 9: 81–88, 2012.

Gerken, M. Internalism and externalism in the epistemology of testimony. *Philosophy and phenomenological research* 87(3): 532–557, 2013.

Gerken, M. Outsourced cognition. *Philosophical Issues* 2014. This volume.

Giere, R. N. Distributed cognition without distributed knowing. *Social epistemology* 21(3): 313–320, 2007.

Gilbert, D. T., D. S. Krull, and P. S. Malone. Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of personality and social psychology* 59(4): 601–613, 1990.

Gilbert, D. T., R. W. Tafarodi, and P. S. Malone. You can't not believe everything you read. *Journal of personality and social psychology* 65(2): 221–233, 1993.

Goldberg, S. C. *Relying on others*. Oxford, 2010.

Goldberg, S. C. Epistemic extendedness, testimony, and the epistemology of instrument-based belief. *Philosophical explorations* 15(2): 181–197, 2012.

Goldman, A. What is justified belief? In: G. Pappas, ed., *Justification and knowledge*, 1–23. Reidel, 1979.

Goldman, A. I. Discrimination and perceptual knowledge. *The Journal of philosophy* 73(20): 771–791, 1976.

Greco, J. The nature of ability and the purpose of knowledge. *Philosophical issues* 17(1): 57–69, 2007.

Green, A. Extending the credit theory of knowledge. *Philosophical explorations* 15(2): 121–132, 2012.

Green, A. Evaluating distributed cognition. *Synthese* 191(1): 79–95, 2014.

Heersmink, R. A taxonomy of cognitive artifacts: Function, information, and categories. *Review of philosophy and psychology* 4(3): 465–481, 2013.

Hirst, W. and G. Echterhoff. Remembering in conversations: The social sharing and reshaping of memories. *Annual review of psychology* 63(1): 55–79, 2012.

Hollan, J., E. Hutchins, and D. Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM trans comput-hum interact* 7(2): 174–196, 2000.

Hutchins, E. How a cockpit remembers its speeds. *Cognitive science* 19(3): 265–288, 1995.

Hutchins, E. *Cognition in the wild*. Bradford, 1996.

Hutchins, E. Enculturating the supersized mind. *Philosophical studies* 152(3): 437–446, 2011.

Johnson, M. K., S. Hashtroudi, and D. S. Lindsay. Source monitoring. *Psychol bull* 114(1): 3–28, 1993.

Kalnikaité, V. and S. Whittaker. Software or wetware? Discovering when and why people use digital prosthetic memory. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, CHI '07, 71–80. New York, NY, USA: ACM, 2007.

Kalnikaité, V. and S. Whittaker. Cueing digital memory: how and why do digital notes help us remember? In: *Proceedings of the 22nd british HCI group annual conference on people and computers: culture, creativity, interaction—volume 1*, BSC-HCI '08, 153–161. British Computer Society, 2008.

Kelp, C. Extended cognition and robust virtue epistemology. *Erkenntnis* 78(2): 245–252, 2013.

Kirchhoff, M. D. Extended cognition and fixed properties: steps to a third-wave version of extended cognition. *Phenomenology and the cognitive sciences* 11(2): 287–308, 2012.

Kirchhoff, M. D. and W. Newsome. Distributed cognitive agency in virtue epistemology. *Philosophical explorations* 15(2): 165–180, 2012.

Koriat, A. The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and cognition* 9(2): 149–171, 2000.

Kornblith, H. *Knowledge and its place in nature*. Oxford, 2002.

Kornblith, H. *On reflection*. Oxford, 2012.

Lackey, J. Why we don't deserve credit for everything we know. *Synthese* 158(3): 345–361, 2007.

Loussouarn, A., D. Gabriel, and J. Proust. Exploring the informational sources of metaperception: The case of change blindness blindness. *Consciousness and cognition* 20(4): 1489–1501, 2011.

Malmgren, A.-S. Review of Relying on Others: An Essay in Epistemology, by Sanford C. Goldberg. *Mind* 120(480): 1251–1258, 2011.

Mascaro, O. and D. Sperber. The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition* 112(3): 367–380, 2009.

Michaelian, K. In defence of gullibility: the epistemology of testimony and the psychology of deception detection. *Synthese* 176(3): 399–427, 2010.

Michaelian, K. Generative memory. *Philosophical psychology* 24(3): 323–342, 2011.

Michaelian, K. Is external memory memory? Biological memory and extended mind. *Consciousness and cognition* 21(3): 1154–1165, 2012a.

Michaelian, K. Metacognition and endorsement. *Mind & Language* 27(3): 284–307, 2012b.

Michaelian, K. (Social) metacognition and (self-)trust. *Review of philosophy and psychology* 3(4): 481–514, 2012c.

Michaelian, K. The information effect: Constructive memory, testimony, and epistemic luck. *Synthese* 190(12): 2429–2456, 2013a.

Michaelian, K. The evolution of testimony: Receiver vigilance, speaker honesty and the reliability of communication. *Episteme* 10(1): 37–59, 2013b.

Michaelian, K. and J. Sutton. Distributed cognition and memory research: History and current directions. *Review of philosophy and psychology* 4(1): 1–24, 2013.

Mitchell, K. J. and M. K. Johnson. Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory? *Psychological bulletin* 135(4): 638–677, 2009.

Oppenheimer, D. M. The secret life of fluency. *Trends in cognitive sciences* 12(6): 237–241, 2008.

Palermos, S. O. Belief-forming processes, extended. *Review of philosophy and psychology* 2(4): 741–765, 2011.

Palermos, S. O. Loops, constitution, and cognitive extension. *Cognitive systems research* 27: 25–41, 2014.

Paynter, C. A., L. M. Reder, and P. D. Kieffaber. Knowing we know before we know: ERP correlates of initial feeling-of-knowing. *Neuropsychologia* 47(3): 796–803, 2009.

Plantinga, A. *Warrant: The current debate*. Oxford, 1993.

Pritchard, D. *Epistemic luck*. Oxford, 2005.

Pritchard, D. Virtue epistemology and epistemic luck, revisited. *Metaphilosophy* 39(1): 66–88, 2008.

Pritchard, D. Cognitive ability and the extended cognition thesis. *Synthese* 175(1): 133–151, 2010.

Pritchard, D. Anti-luck virtue epistemology. *Journal of philosophy* 109(3): 247–279, 2012.

Proust, J. *The philosophy of metacognition: Mental agency and self-awareness*. Oxford University Press, 2013.

Rajaram, S. and L. P. Pereira-Pasarin. Collaborative memory: Cognitive research and theory. *Perspectives on psychological science* 5(6): 649–663, 2010.

Reber, R. and C. Unkelbach. The epistemic status of processing fluency as source for judgments of truth. *Review of philosophy and psychology* 1(4): 563–581, 2010.

Reder, L. M. and F. E. Ritter. What determines initial feeling of knowing? familiarity with question terms, not with the answer. *Journal of experimental psychology: Learning, memory, and cognition* 18: 435–451, 1992.

Ren, Y. and L. Argote. Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *The Academy of management annals* 5(1): 189–229, 2011.

Robbins, P. and M. Aydede, eds. *The Cambridge handbook of situated cognition*. Cambridge, 2008.

Roediger, H. L., M. L. Meade, and E. T. Bergman. Social contagion of memory. *Psychonomic bulletin & review* 8(2): 365–371, 2001.

Roth, W.-M. and A. Jornet. Situated cognition. *WIREs Cogn Sci* 2013.

Rupert, R. *Cognitive systems and the extended mind*. Oxford, 2009.

Rupert, R. D. Memory, natural kinds, and cognitive extension; or, martians don't remember, and cognitive science is not about cognition. *Review of Philosophy and Psychology* 4(1): 25–47, 2013.

Smart, P. Extended memory, the extended mind, and the nature of technology-mediated memory enhancement. In: *1st ITA Workshop on Network-Enabled Cognition: The Contribution of Social and Technological Networks to Human Cognition*. 2010.

Sosa, E. *A virtue epistemology: Apt belief and reflective knowledge*, vol. 1. Oxford, 2007.

Sparrow, B., J. Liu, and D. M. Wegner. Google effects on memory: Cognitive consequences of having information at our fingertips. *Science* 333(6043): 776–778, 2011.

Sperber, D. Speakers are honest because hearers are vigilant: Reply to Kourken Michaelian. *Episteme* 10(1): 61–71, 2013.

Sperber, D., F. Clément, C. Heintz, O. Mascaro, H. Mercier, G. Origgi, and D. Wilson. Epistemic vigilance. *Mind & Language* 25(4): 359–393, 2010.

Sterelny, K. Minds: extended or scaffolded? *Phenomenology and the Cognitive Sciences* 9(4): 465–481, 2010.

Stokes, D. Cognitive penetrability of perception. *Philosophy Compass* 8(7): 646–663, 2013.

Stone, C. B., A. Coman, A. D. Brown, J. Koppel, and W. Hirst. Toward a science of silence: The consequences of leaving a memory unsaid. *Perspectives on Psychological Science* 7(1): 39–53, 2012.

Sutton, J. Distributed cognition: Domains and dimensions. *Pragmatics & Cognition* 235–247, 2006.

Sutton, J. Exograms and interdisciplinarity: history, the extended mind, and the civilizing process. In: R. Menary, ed., *The extended mind*, 189–225. MIT, 2010.

Sutton, J., C. Harris, P. Keil, and A. Barnier. The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences* 9(4): 521–560, 2010.

Theiner, G. Transactive memory systems: A mechanistic analysis of emergent group memory. *Review of Philosophy and Psychology* 4(1): 65–89, 2013.

Vaesen, K. Knowledge without credit, exhibit 4: extended cognition. *Synthese* 181(3): 515–529, 2011.

Vrij, A. *Detecting lies and deceit: Pitfalls and opportunities*. Wiley, second edn., 2008.

Weldon, D. M. Remembering as a social process. In: D. Medin, ed., *Psychology of learning and motivation: Advances in research and theory*, 67–120. Academic, 2001.

Yeo, R. Notebooks as memory aids: Precepts and practices in early modern England. *Memory Studies* 1(1): 115–136, 2008.

Zagzebski, L. *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge, 1996.