



Believing on Trust

Kappel, Klemens

Published in:

Synthese - An international journal for epistemology, methodology and philosophy of science

DOI:

[10.1007/s11229-013-0376-z](https://doi.org/10.1007/s11229-013-0376-z)

Publication date:

2014

Document version

Early version, also known as pre-print

Citation for published version (APA):

Kappel, K. (2014). Believing on Trust. *Synthese - An international journal for epistemology, methodology and philosophy of science*, 191(9), 2009-2028. <https://doi.org/10.1007/s11229-013-0376-z>

Believing on Trust[#]

by Klemens Kappel, Division of Philosophy, University of Copenhagen, Denmark, kappel@hum.ku.dk

Published in Synthese (2014) 191:2009–2028

DOI 10.1007/s11229-013-0376-z

Abstract

The aim of the paper is to propose a way in which believing on trust can ground doxastic justification and knowledge. My focus will be the notion of trust that plays the role depicted by such cases as concerned John Hardwig in his early papers (Hardwig 1985; Hardwig 1991), papers that are often referenced in recent debates in social epistemology. My primary aim is not exegetical, but since it sometimes not so clear what Hardwig's claims are, I offer some remarks of interpretation that might be of interest. The main purpose of the paper, however, is this: following various cues in Hardwig's writing, I specify certain epistemic properties of agents in social systems, such that, roughly speaking, for agents to know (or be justified in believing) what the 'system knows', social relations of epistemic trust between agents in the system are necessary. I will

* Research on this paper has been generously funded by the Velux Foundation. Earlier versions of this material was presented at a departmental seminar, a SERG Workshop on epistemic trust, December 2010, the 3rd Copenhagen Conference in Epistemology, August 15-17 2011, and 2nd Berlin Conference on Meta-Epistemology, September 1-3 2011. I thank participants at these events for patience and helpful comments [in alphabetical order]: Kristoffer Ahlström, Finn Collin, Paul Faulkner, Axel Gelfert, Mikkel Gerken, Sandy Goldberg, Kathrine Hawley, Chris Kelp, Arnon Keren, Jennifer Lackey, Brent Madison, Kirk Michelian, Martin Kusch, Duncan Pritchard, Gloria Orrigi, Ernest Sosa, Nikolaj Jang Lee Linding Pedersen, Shane Ryan, Sebastian Rödl, Chris Thompson, Michael Williams, Frank Zenker.

suggest that we can view this social form of epistemic trust as non-inferential dispositions to believe what some individual or other source of information asserts or transmits. When this disposition is discriminating and defeater-sensitive, it can ground knowledge and justification. Or, more cautiously, we should be sympathetic to this view if we are inclined to accept the core insight of process reliabilism. Finally, I will offer some remarks about how epistemic trust and epistemic reasons may relate on this picture.

1. Believing on trust

The concern in the following is believing on trust, that is, believing a proposition because one trusts some person, institution or indeed tradition conveying it. The aim of the paper is to propose a way in which believing on trust can ground epistemic justification and knowledge. The main examples I will be concerned about are those provided by Hardwig, where he claims that beliefs accepted upon trust among members of the scientific community is essentially involved in much of what we normally regard as scientific knowledge (Hardwig 1985; Hardwig 1991).

I will assume that in cases like these we accept certain beliefs as true, or likely to be true, because we trust a human or institutional source asserting the proposition in question. So in a sense, trust is our reason for accepting the belief. And we often express this by saying things like: 'I believe that p because my friend told me, and I trust my friend on these matters'. Often, of course, trust in a source will only be one part of the full explanation of why one adopts a particular belief. So, it might typically be more accurate if also more cumbersome to say that we accept a belief in part because we trust some source.

I will use the label 'epistemic trust' when I talk about trust involved in these and similar cases where we adopt beliefs on the basis of trust. The overall aim of the paper is to outline a theory of epistemic trust. More specifically, I want to propose a process reliabilist account of how we can acquire knowledge and justified

belief on grounds of epistemic trust. As we will see, the project bears many similarities to the programme of (Goldberg 2010), from which I have learned a lot. There are other prominent uses of the word 'trust', or 'rational trust', and I will comment briefly on some below, but my main concern is not to discuss these.

It is natural for philosophy to ask what epistemic trust is. What is the nature of the attitude or the relation? Does epistemic trust in *A*, for example, reduce to having certain *beliefs* about *A*? If so, epistemic trust would be a doxastic attitude, and accepting beliefs on grounds of trust would be a special case of something quite familiar in epistemology. One alternative to this picture would be to view epistemic trust as distinct kind of non-doxastic attitude, say something like a feeling of confidence or an emotion. If this is how we think of epistemic trust, then it seems that any significant epistemic role of trust would require some explanation. Below, I shall propose that we view epistemic trust as neither a doxastic attitude, nor as a non-doxastic psychological attitude. We can, I want to suggest, view epistemic trust as a dispositional property.

Another interesting question concerns whether or how epistemic trust can ground knowledge and justification? Can epistemic trust have reasons? Can epistemic trust without reasons be rational? Can epistemic trust interact with reasons and evidence, and if so how?

Before trying to address such questions, we should, I believe, note an important caveat. Generally, we should not presuppose that trust is a distinct uniform type of psychological attitude or interpersonal relation. The word 'trust' in ordinary parlance does not denote a unified mental or social kind, or at least we cannot take this for granted at the outset of our philosophical theorising. Rather, there seems to be a fairly heterogeneous family of related phenomena that may be variously picked out when we talk about trust, or this is what I suspect at any rate. Something similar holds, of course, of the label 'epistemic trust' except that this phrase is not even part of ordinary parlance. It is a term of art that does not have any interesting meaning apart from the one we invest in it. In

consequence, different accounts of trust in the philosophical literature may very likely conceptualize different phenomena going under the same name.

In view of this fact, my strategy will be the following. Though my aim is not exegetical, I will focus on the kind of trust that figures in John Hardwig's well-known papers (Hardwig 1985; Hardwig 1991). Hardwig argued that the sort of pervasive division of cognitive labour now prevalent in the sciences requires extensive believing on trust (more on this below). My focus will be the notion of trust that plays the role depicted by such cases as concern Hardwig. There are, to be sure, other notions of epistemic trust in the literature, but I shall have little to say about them.

Thus, I will proceed as follows. Following various cues in Hardwig's papers, I specify certain epistemic properties of agents in social systems, such that, roughly speaking, for agents to know (or be justified in believing) what the 'system knows', relations of epistemic trust between agents in the system are necessary. So what I propose is, in a sense, a functional account of epistemic trust, as distinct from for example a genealogical account of epistemic trust (Faulkner 2007; Kappel 2010). I will suggest that we can view epistemic trust as non-inferential disposition to believe what some individual or other source of information asserts or transmits. When this disposition is discriminating and defeater-sensitive, it can ground knowledge and justification. Thirdly, I will explain how, given standard process reliabilism, beliefs adopted on this basis trust can indeed qualify as known and justified belief. Finally, I will offer some remarks about how epistemic trust and epistemic reasons may relate on this picture.

2. Trust as socio-epistemic glue

In his well known papers on trust and knowledge (Hardwig 1985; Hardwig 1991), Hardwig observed that scientific research commonly is the product of collaboration involving many individual researchers, indeed sometimes hundreds of researchers, each producing sub-results within their own specialties that provide the individual pieces that in concert substantiate the target finding. He

illustrates this by cases drawn from experimental physics and mathematics, but something similar is no doubt true of research in many areas. This combination of collaboration and specialisation has, Hardwig suggests, a crucial epistemic implication: members of the collaborating group often have no choice but to trust the scientific competence and integrity of their associates for the accuracy of their results, and the honesty of their reports. This, I will assume, is the role of epistemic trust.

These examples and others like them concern what might be called *specialised collaborative epistemic projects*. But this I roughly mean projects that aim to establish the truth (or falsity) of a set of target propositions by means of collecting a large set of data and performing a number of inferences, where data collection and inferences is partitioned between several of individuals, each of whom is a specialist in his or her own particular field, but not in all fields involved in the project. Accordingly, in specialised collaborative projects each part of the story might be known by someone, and yet there may be no single individual who knows every part of the story. When this is so, the obvious epistemological question we can ask is this: Does anyone know the truth of the target propositions (assuming that they are true)? And if the target proposition is known, then by whom is it known, and how?

As indicated, Hardwig's basic proposal is this. In specialised collaborative epistemic projects, epistemic trust might (with other conditions) be sufficient for the individual members to know or be justified in believing the conclusions of such projects. And while there may be other conditions that would also be sufficient, epistemic trust is often the only relation that will actually be in place. Thus, epistemic trust is often what actually enables members of specialised collaborative epistemic projects to know or justifiably believe the conclusions of these projects.

2.1 A puzzling first account

To arrive at a more detailed account of the functional role of epistemic trust, consider first this somewhat schematic way of

characterising the epistemic properties of specialised collaborative epistemic projects. Suppose we have a socio-epistemic structure S consisting of the following

Premises P_1-P_n , a conclusion C , evidential relations between premises and the conclusion C , evidence for each premise $E_1 - E_n$, a group of agents A_1-A_n , A_n ,

for which the following holds:

(1) For each premise P_i , exactly one agent A_i knows (or justifiably believes) the premise by first hand possession of the relevant evidence E_i .

(2) Exactly one agent A_n understands the evidential relationship between P_1-P_n and C , so only this agent knows (or justifiably believes) that the evidential relationship between P_1-P_n and the conclusion holds.

(3) All agents knows (or justifiably believe) that (1) and (2) is true of S .

Are (1), (2) and (3) compatible with any agent in S knowing the conclusion C ? One might wonder how agents could know C . After all, every agent seems to lack knowledge or evidence crucial for inferring C ? As I said, the Harwigian response I am interested in points to some form of social trust. In general, agents can know or justifiably believe conclusions of specialised collaborative epistemic projects partly in virtue of trusting that the other members of the collaborative project have indeed succeeded in doing their parts.

However, before elaborating this, I briefly want to comment on what turns out to be an entirely different suggestion also aired by Hardwig. Hardwig invites us to consider (and accept) the following principle (Hardwig 1991, 697):

(T) If A has good reason to believe that B has good reasons to believe p, then A has good reasons to believe p.

and the similar principle for knowledge:

(T') If A knows that B knows p, then A knows p.

Presumably, Hardwig's point is that if we grant these principles, we can explain how agents in a socio-epistemic structure like S may know or be epistemically justified in believing C, despite not initially knowing or being epistemically justified in accepting the epistemic foundation of C. This is so, of course, in virtue of (3) above, in conjunction with (T) and (T'). When these assumptions are combined with assumption (3), then we get, roughly, that what any agent in S knows, all agents in S know. And we get that what any agent in S has good reason to believe, every agent has good reason to believe. So, (T) and (T') would explain how agents in S know or justifiably believe C.

Interestingly, Hardwig seems to think that there is a distinctive 'blindness' to the knowledge that each agent might acquire in this way. The reasons necessary for the agent's knowledge may include reasons that the agent does not have, but only knows or justifiably believes that other agents have (Hardwig 1991, 699). Yet, despite this blindness, they know, or so Hardwig suggests.

Indeed, it might be worries about this very blindness that leads Hardwig to think that (T) and (T') are controversial and therefore in need of justification. At any rate, Hardwig seems concerned about why we would want to accept (T) or (T'), and he proposes an argument in favour of them. The argument Hardwig proposes has the form of a dilemma. Either we accept (T) and (T'), and the ensuing blindness of knowledge. Or, and this is the other horn of the dilemma, we face the following trilemma (Hardwig 1991, 699):

(i) Skepticism. There can be no knowledge or justified belief in scientific disciplines in which there is 'too much available evidence'.

(ii) Irrelevance of evidence. We can have knowledge while being ignorant of evidence: 'one can know p only by ignoring most of the best evidence', or at least despite ignoring the best evidence.

(iii) Epistemic collectivism. Knowledge and justified belief can pertain to collectives, rather than to individuals.

Hardwig rejects the first two options in this trilemma as implausible, and assumes for the sake of further investigation that we ignore (iii). But then we seem forced to accept (T) and (T'). Now, there are some issues of interpretation regarding the details of the argument, and the way that Hardwig sets out the trilemma. But for the purpose of the discussion below I want to set these questions aside. The main point that Hardwig seems intent to make is that we need to accept (T) and (T') on pain of being unable to explain how one might know or have justified beliefs in the conclusions of collaborative epistemic projects.

Actually, what Hardwig says in defence of (T) and (T') is puzzling for at least two reasons. First, note that if we accept.

Factivity. S knows that $p \Rightarrow p$

and

Closure. S knows that p & S knows that $(p \Rightarrow q) \Rightarrow S$ knows that q

then (ignoring a range of qualifications that hopefully don't matter here) it follows that

(T') If A knows that B knows that p, then A knows that p.

So, (T') as stated may not be controversial at all, since it is a trivial implication of principles that are widely accepted. This would provide a defence of (T') that is much more persuasive than the sort of *reductio* that Hardwig offers. Whether (T) is controversial very much depends on the details of what is meant by 'good reasons', but it seems quite tempting to think that many plausible views of justification will support principles such as (T). So, contrary to what Hardwig seems to imply, (T) and (T') may not be controversial, and need not be defended by the sort of reduction that Hardwig presses.

The second and perhaps more puzzling feature of this part of Hardwig's paper is that it is not obvious what all of this has to do with epistemic trust. A's having *reason* to believe that B has a good reason to believe that p, or A's *knowing* that B has good reason to believe that p, does not invite to saying that A trusts B on this matter. Thus, if (T) and (T') are the principles by which structures such as S are held together, then one might reasonably say that they are linked by chains of justified belief, or chains of knowledge, rather than by trust. It is by no means obvious that what enables members of S to know the conclusion S has anything to do with what we might reasonably call epistemic trust.¹

¹ The impression that trust has dropped out of the picture is also conveyed by Hardwig's own summary of the transmission of knowledge in social systems (Hardwig 1991, 699):

- (1) A knows that B says that p.
- (2) A believes (and has good reasons to believe?) that B is speaking truthfully, i.e., that B is saying what she believes.
- (3) A believes (and has good reasons to believe?) that B (unlike A) is in a position, first, to know what would be good reasons to believe p and, second, to have the needed reasons.

2.2 Trust without evidence

I want to suggest that the more interesting question we should consider is the following. Unlike what we stipulated above, in specialised collaborative epistemic projects such as those common in science, members of these projects will *not* be in a position to know or be justified in believing that other members really have good reasons or evidence for believing what they do about their part of the project. That is, (3) or its equivalents in the specification of S above will not be true. When this is so, principles (T) and (T') do not apply. So, should we say that such projects do not enable their members to know or justifiably believe the conclusions of their projects?

Again, to make this question more transparent, consider a new schematic structure S', consisting of

Premises P_1 - P_n , a conclusion C, evidential relations between premises and the conclusion C, evidence for each premise E_1 - E_n , a group of agents A_1 - A_n , A_n ,

for which the following holds:

(1) For each premise P_i , exactly one agent A_i knows or justifiably believes the premise by first hand possession of the relevant evidence E_i

(2) Exactly one agent A_c understands the evidential relationship between P_1 - P_n and C, so only this agent knows or

(4) A believes (and has good reasons to believe?) that B actually has good reasons for believing p when she thinks she does.

Again, no mention of trust or epistemic trust is made in these principles. Note also the curious fact that the *antecedents* in (2), (3), (4) do *not* correspond to (T), though the paranteses do (ignoring the question marks).

justifiably believes that the evidential relationship between P_1 - P_n and the conclusion holds.

(3') No agent has sufficient evidence for knowing or justifiably believing that (1) and (2) holds of S' , though they do not have evidence suggesting that (1) and (2) are false.²

Clearly, (T) and (T') do not apply to S' , and this is because (3') states that no agent in S' is justified in believing (1) and (2), or knows that these hold of S' . That is, for all the agents know or justifiably believe, there might be some other members of the collective epistemic project who are not in possession of the appropriate first hand evidence, or do not justifiably believe or know the premises that they were supposed to be in charge of.

I actually think that this is the kind of socio-epistemic system that Hardwig really meant to discuss. At least, this interpretation makes sense of Hardwig's elaborate discussion of two further key questions. First, is science often like S' ? And second, are agents in systems like S' in a position to know the conclusion C ? Clearly, Hardwig defends an affirmative answer to both questions.³

² For consistency reasons, we can make an exception for A_c , such that (3') reads: with the exception of A_c , no other agent knows or justifiably believe that A_c is justified in believing that the proper evidential relationship between P_1 - P_n and the conclusion C actually holds.

³ Hardwig defends an affirmative answer to the first question roughly as follows. Say that an agent A is trustworthy with respect to some domain of facts D just in case the agent (on request) reports her beliefs regarding D in a truthful, competent, and conscientious manner (Hardwig 1991, 700). Generally speaking, then, if an agent A is *trustworthy*, then if A reports that p , A is justified in her belief that p , or knows that p . In science, could one know or have first hand justified beliefs that some other researcher in some entirely different speciality than one's own, is trustworthy? Often the answer is 'No'

Here I want to restrict myself to the second question just cited. How, if at all, can any member of S' know the conclusion C? Here is how Hardwig sums up what appears to be the bottom line:

'a scientific community has no alternative to trust, including trust in the character of its members.' (Hardwig 1991, 706).

Or here in a more elaborate form:

'A must TRUST B, or A will not believe that B's testimony gives her good reasons to believe p. And B must be trustworthy or B's testimony will not in fact give A good reasons to believe p, regardless of what she might believe about B. A team of scientific experimentalists, for example, must both trust each other and be worthy of that trust or their experiment will not give anyone enough good reason to believe their conclusions. ' (Hardwig 1991, 700)

(Hardwig 1991, 701-2). Hardwig considers and rejects some reasons that might suggest otherwise. Does game-theoretic reasons show that individual researchers have prudential reasons to become and remain trustworthy? Often the answer is 'No'. Defection from cooperation may be detected too rarely, and there are too few interactions (Hardwig 1991, 702 ff.). Are there deterrents or safeguards in the regulation of science that would effectively detect or prevent fraudulent characters, thus grounding an assumption that scientists are generally trustworthy? It seems not (Hardwig 1991, 703-5). So, the conclusion is that scientific communities (or sets of such communities) might often be social systems like S'. Although researchers may believe that other researchers are trustworthy, they often do not have first hand justification for such beliefs, and they do not know that others are truthsworthy. At best, this is just something they merely believe, or hope.

I suggest that the basic idea proposed here might be captured roughly in the following way:

Trustworthiness. An agent A is *trustworthy* with respect to some domain of facts D just in case A is (i) *competent* with respect to questions within this domain, i.e. the agent can decide questions within this domain with a rather high reliability, and (ii) *sincere*, i.e. on request or when otherwise relevant for the audience, A honestly reports her beliefs regarding the truth of propositions in D.

In general, then, when people are trustworthy with respect to a particular domain, they are likely to have true beliefs about questions in that domain, and they will be honest and forthcoming in their reports about the content of these beliefs. I do not suggest, of course, that this definition of trustworthiness is precise, but it is precise enough for the discussion here. I want to propose that the basic Hardwigian suggestion is that what provides the epistemic glue of systems like S' is not principles like (T) or (T'), but rather a principle like the following:

(T*) If A knows that B reports that p, where p is in domain D, and A trusts B with respect to domain D, and B is in fact trustworthy regarding domain D, and A thereby adopts the belief that p, then A knows that p, or is epistemically justified in her belief that p.

There are, of course, many further questions that one might press regarding (T*), including what the epistemic status of B's belief in p must be for A to get to know or be epistemically justified in believing that p. I want to set aside these questions here. Our interest in (T*) is due to the fact that this principle identifies a role for trust in the interpersonal transmission (or generation) of knowledge and justification. Even if A does not know and does not justifiably believe that B is trustworthy with respect to D, then A may trust that this is

so. And, according to (T*), when B is trustworthy, A's trust that this is so might be enough for A's resulting beliefs to earn a significant epistemic status.

3. A Process Reliabilist Underpinning of Epistemic Trust

Return now to our main question: What is the nature of epistemic trust? We can now try to avoid the essentialist caveat noted earlier by asking this question as follows: What is the nature of the relation that might suit the role designated by 'A trusts B' in (T*), such that (T*) comes out as a plausible epistemic principle, explaining how members of S' can know or be epistemically justified in the conclusion C?

It is worth here recalling two suggestions. First, one might think that A trusts B just in case A knows or justifiably believes that B is trustworthy with respect to a particular domain. While surely not unreasonable in its own right, the problem in this context is that on doxastic notions of trust, the configuration of S' and (T*) would imply that members of S' do not know the conclusion C, or are not epistemically justified in this conclusion, since we have assumed that in such systems agents do not know, and are not justified in believing, that other agents are trustworthy within their specialised domains.⁴

Alternatively, one might think of trust as a particular emotion or a feeling of confidence about someone or something, which would likely be a non-doxastic psychological state. Again, considered in its

⁴ As a reviewer pointed out, one could imagine other doxastic accounts of epistemic trust. One might say, for example, that A trusts B just in case A believes that B is trustworthy, though A's belief need have no special pedigree, and therefore no special epistemic status. As I have stipulated S', nothing rules out that agents could trust each other in this slim sense. The challenge for this form trust is how it would explain that it can ground knowledge (assuming that we want our account of trust to explain this). Presumably, more conditions would need to be added.

own right, this idea does not seem implausible. But for such a notion of trust to fit into (T*) much more would have to be said.

For these reasons I want to explore the possibility that we think of trust along the following lines:

A trusts B regarding a domain D just in case A has a non-doxastic disposition to accept as true (or give high credence to) the content of B's statements within domain D.

A non-doxastic disposition is here a disposition that is not itself based on a belief that someone is trustworthy. So, on this view, epistemic trust is neither a form of belief, and not is it a non-doxastic attitude akin to hopes, emotions or desires.

Let me hasten to point out that the notion of trust just given will not alone render (T*) plausible. This is for the following reason. Suppose that A trusts B who is in fact trustworthy, but A's trust in B is quite indiscriminate - A would trust just about anyone, and it was mere luck that made A trust someone who is in fact trustworthy. Or suppose that A trusts B, but that A ignores evidence suggesting that B is not trustworthy at all in the relevant domain. In both types of cases we would, I submit, be disinclined to accept that A's resulting beliefs would be justified, or count as known if true. So, the proposed notion of trust does not meet the requirement of rendering (T*) plausible, and this is because a condition on discrimination in the selection of trusted sources, and a condition on sensitivity to defeaters is not met. One might consider building these conditions into a somewhat artificial notion of trust, which, when appended to (T*) would have this principle come out plausible. But rather than doing that, it seems more transparent and natural simply to revise (T*) to get the following:⁵

⁵ Thanks to Mikkel Gerken for suggesting that this is a more natural strategy.

(T**) If A knows that B reports that p, where p is in domain D, and A trusts B with respect to domain D, and (i) B is in fact trustworthy regarding domain D, and (ii) A's trust in B is discriminating and defeater-sensitive, and (iii) A thereby adopts the belief that p, then A knows that p, or is epistemically justified in her belief that p.

A disposition to accept as true is discriminating just when it tends to be pick out trustworthy sources. A trusting agent has the property of defeater-sensitivity when she remains sensitive to cues or evidence that the information source is not trustworthy, perhaps despite her initial trust in the source.

Let me add some remarks about this characterisation of epistemic trust. It might be tempting to think of trust as a distinct type of psychological attitude or emotion that one might take to other individuals. However, my being inclined to accept what you say as true regarding some matter is not the same as entertaining any particular psychological attitude to you. But epistemic trust thus conceived may be an emotionally based disposition, by which I mean that it may be in part caused or upheld by an emotional attitude. I may epistemically trust someone where this trust is causally dependent on my liking this person, or my feeling confident about the person. I may epistemically trust someone because I am in the habit of doing so, or because doing so is an inseparable part of a membership of a cultural group, or indeed a research tradition. So, epistemic trust may be causally related to psychological or sociological mechanisms not traditionally counted as a part of epistemology.

As mentioned, (T**) contains a condition regarding discrimination. This is motivated by the fact that acquiring knowledge by trusting a source of information is not compatible with trusting everything or everyone, even if one by chance stumbles upon a trustworthy source. So, for true beliefs acquired by trust in sources to acquire the status as knowledge, these sources need to be selected in certain discriminating ways. So, knowledge based on

trust not only requires that one's trusted source is in fact reliable, but also that the mechanism by which one has selected a particular source is non-arbitrary in certain ways. But the mechanisms that facilitates the discrimination need not, on the view I am proposing, be internal to the trusting subject, and he or she need not even know about their existence, or justifiably believe that they operate. A trusting subject may be embedded in a social system that brings about the discrimination necessary, and I suspect that this is often how we manage acquire sufficiently discriminatory power.

Knowledge or justification acquired by epistemic trust also requires sensitivity to defeaters. Someone placing epistemic trust in some source of evidence must remain sensitive to cues or evidence that the source is not trustworthy after all. An agent might for example acquire evidence suggesting that a source is in general not reliable (general undermining defeater), or that an otherwise trustworthy subject has made a mistake on a particular occasion (specific undermining defeater), or against the proposition asserted by the trusted person (specific overriding defeaters). I do not want to suggest that any defeating evidence, however weak or remotely plausible, can undermine conditions under which we can acquire knowledge or justification by epistemic trust in a source. The requirement of sensitivity that I have in mind is only the unspecific one that when presented with strong enough defeaters, the trusting subject should not accept the belief or information relayed to her.

Finally, one may wonder why condition (i) and (ii) are both necessary.⁶ The thought would be that (ii) makes (i) redundant. If A's trust in B is genuinely discriminating and defeater-sensitive, then B has to be trustworthy, so we do not need to add (i) as a separate requirement. However, Suppose that some agent B is not trustworthy (meaning either that B is not competent, or that B is not inclined to transmit what B takes to be true). Some agent A can be sensitive to defeaters, and in possession of a (general) capacity to discriminate trustworthiness, and yet A trusts B (since no defeaters

⁶ Thanks to an anonymous reviewer for rasing this worry.

to B's trustworthiness are around, and because A's capacity to discriminate trustworthiness is not infallible. So, A has mistakenly identified B as trustworthy. In this case, even if A happens to get a true belief from B, this is not knowledge. It's Gettierized knowledge, if you will. So, condition (i) is needed.

Of course, I don't claim that whenever we trust some source, we do so in way that is discriminate and defeater-sensitive. If someone is disposed to accept as true what various sources says, but his trust does not meet the requirement of discrimination, we can refer to this as *indiscriminate trust*. True belief acquired by indiscriminate trust does not amount to knowledge. Similarly, if someone trusts some source of information, and yet remain insensitive to information defeating the trustworthiness of the source, then we have a case of what we might call *defeater-insensitive trust*. Again, defeater-insensitive trust cannot ground knowledge. Both are forms of gullibility, and both might be considered marks of epistemic irrationality.⁷

3.1 Trust based super-credence

It is crucial to consider whether we ever believe much on epistemic trust as I have stipulated this notion. As we have seen, Hardwig seems to think that we do, indeed he insists that the vast majority of what we regard as scientific knowledge depend on trust, and as I discussed above, arguably what he had in mind was something like epistemic trust in the stipulated sense.

In part, Hardwig's argument for the central role of trust depend on a negative assessment of the possibility that members of the scientific community could know or be justified in believing that collaborating partners in the scientific community are indeed trustworthy in the sense presupposed in such collaborative epistemic projects. However, it has been suggested that Hardwig is too pessimistic here, underestimating the extent to which members of various scientific specialties might have good evidence for the

⁷ See the related discussion in (Goldberg 2010)

trustworthiness of other parts of science, see (Adler 1994; Goldman 2001).

Another challenge comes from the view that there are general but defeasible reasons to expect that others are trustworthy in the domains that they speak out about: when someone asserts something, then one has a *prima facie* reason to believe the truth of what has been asserted. If this is right then while we might well epistemically trust others, the crucial epistemic work could be due these sorts of *prima facie* reasons to expect that that others are trustworthy (Burge 1993; Foley 2001).

Yet at third question concerns whether we could reasonably attribute epistemic trust to an agent without attributing the corresponding beliefs about trustworthiness. The idea is this: if we think that A is disposed to accept as true what B reports in a particular domain, then it would be quite natural also to attribute to A the *belief* that B is trustworthy in this domain. But now it may begin to look as if, when we have good reasons to say that A epistemically trusts B, we also have good reasons to think that A believes with at least some confidence that B is trustworthy. But it may be hard to attribute such a belief to A, unless we also think that A has at least some evidence for the trustworthiness of B. But now it seems that what does the epistemic work need not be A's epistemic trust in B as such, but A's associated beliefs and evidence regarding the trustworthiness of B.

For such reasons, one might worry about the idea that we often accept beliefs on grounds of epistemic trust, and we might even doubt whether this is a coherent possibility. At least it should be stressed that this is not nearly as obvious as might seem at first impression. When relations of epistemic trust are in place, many other epistemically relevant factors may also be operative.

Nonetheless epistemic trust remains an interesting topic for epistemology, or so I want to maintain. To focus our discussion, we can consider cases in which someone is more confident in her belief than is warranted by the evidence that she has for the truth of the belief, and where this higher level of confidence is due to his

epistemic trust in some source of information. More specifically, we can consider what I will label *trust-based super-credence*. Cases of trust-based super-credence are instances of the following schema:

- (i) A believes that p with credence n
- (ii) A's own evidence for p warrants credence m
- (iii) $n > m$
- (iv) A's super-credence is due to A's epistemic trust in source I (where I is a source informing A that p).

Here are some cases that illustrate ways in which trust-based super-credence might come about. My friend Paul tells me that p. When I consult my evidence, it turns out that I only have some rather weak evidence that Paul is trustworthy in this domain, deriving from my previous dealings with Paul, and from my general assumptions about Paul's epistemic competence in the domain in question. However, I also have some evidence (again, rather weak evidence) suggesting that p is false. Nonetheless, because I epistemically trust Paul in these matters, I believe that p with higher credence than my own evidence warrants.

Another case is this. My friend Peter tells me that p. Though evidence about the trustworthiness of Peter is available to me, I don't consider it and I don't process it. I do not even consider whether such evidence is available - I neither believe that it is available or believe that it is not. I just epistemically trust Peter on the matter, and thus believe that p with a higher credence than is warranted by the evidence that I in fact consider.

Here is a third case. My friend Oscar tells me that p. I am aware of some evidence for p, and of some evidence for the trustworthiness of Oscar in these matters. However, I do not try to determine the exact degree to which these sources of evidence support p. Instead, because I have considerable epistemic trust in

Oscar on these matters, I place a higher credence in the truth of p than is actually warranted by my reflection on the evidence.

Cases of trust-based super-credence might be rare, though I am inclined to think they are not. However that may be, for the reasons I have provided, they seem possible. What is much more debatable, of course, is whether they are signs of epistemic irrationality. Or more specifically, one might suspect that putative cases of trust-based super-credence face a dilemma. Either these cases are incompletely described - there is bound to be some hidden evidence in these cases, evidence that rationalize the level of credence adopted by the trusting agent. Or else they are cases of epistemic irrationality. The first horn of this dilemma asserts that one cannot really have higher credence in a belief than one's evidence warrants - real cases of this sort inevitably contain additional strains of evidence concerning the trustworthiness of the source in question. The second horn of the dilemma admits that this need not be so, but charges that we then face a case of epistemic irrationality.

3.2 A process reliabilist account of epistemic trust

I want to adapt a well-known albeit philosophically controversial reply to this problem. We can make sense of the suggested notion of epistemic trust on the wider background of process reliabilism regarding justification (Goldman 1979; Goldman 1986). Essentially, this move evades the above dilemma by simply denying that justified belief or knowledge needs to be based on evidence (in the sense of evidence necessary for the above dilemma to arise). So, according to process reliabilism, there is nothing problematic about according a belief high credence when doing so is due to a reliable belief-forming process, and there are no defeaters in the vicinity. I will not defend process reliabilism or related views; there are all sorts of well-known worries about process reliabilism. My much more modest aim is merely to elaborate the idea that epistemic trust can be considered an ingredient in belief forming process that may yield knowledge in much the same way that memory or perception do, according to the process reliabilist. Also, note that I do not want to

claim that process reliabilism uniquely supports a notion of epistemic trust - other externalist views may well do so for reasons parallel to the ones depicted below.⁸

On the view I propose, acquiring a belief by epistemic trust in a trustworthy source may yield doxastic justification, or knowledge if the belief is true, and the level of doxastic justification high enough. What makes a belief acquired by trust known or justified is the nature of the trusting disposition in conjunction with the trustworthiness of the trusted source, and the conditions on discrimination and defeater-sensitivity. According to the process reliabilist picture I adopt, the explanation of this is due to the internal relation between between epistemic trust in trustworthy sources and reliable belief-forming processes. Very roughly, when a subject acquires a belief by epistemic trust in a trustworthy source, this implies the existence of a reliable belief-forming process by which the belief is acquired or sustained. And when the conditions on discrimination and defeater-sensitivity are met, the reliable process in question has been selected in discriminating ways, and the trusting subject is sensitive to defeaters to the reliability of that process. This is the process reliabilist explanation of why epistemic trust can generate knowledge and justified belief.

On the view emerging, epistemic trust in external sources of information is analogous to the way that we trust our senses, our memory and various inferential skills. According to a rough process reliabilist account, these are all ingrained belief-forming processes that may ground knowledge and doxastic justification in so far as they work reliably, and provided that the processes meet the conditions of discrimination and sensitivity. The condition regarding defeater-sensitivity is normally met, as we are usually sensitive to defeating evidence regarding the reliability of our modes of sense perception, memory and inferential skills. We can imagine a scenario in which a being uses these native reliable modes of belief formation,

⁸ This parallels the strategy in (Goldberg 2010), to which my exposition here owes a lot.

but where the subject nonetheless ignores cues suggesting that the mode is unreliable. Where the evidence of unreliability is strong enough, this is incompatible with knowledge, as well as with doxastic justification.

We can imagine a being that fails to be discriminate regarding her use of native cognitive resources. Imagine a cognitive being endowed with several native or acquired cognitive resources, say several modes of sense perception, or patterns of reasoning. Suppose that for some reason only half of these processes are reliable, given the tasks they are applied to. Suppose now that the cognitive being has the capacity to trust any mode of belief formation, but that this trust is indiscriminate. So, for a given cognitive task, the being happens to trust certain modes of belief formation, but the selection of modes is completely indiscriminate between reliable and unreliable sources. Now, intuitively, a belief formed in this way would not count as epistemically justified, or as known if true. This is because the condition on discrimination is not met. Fortunately, we are not cognitive beings like that. Evolution, reflection and acquired skills usually make us highly discriminate in the selection of the cognitive sources we trust for a given task.

In many of the cases we have considered, epistemic trust is directed at individuals, groups of individuals, institutions, or even traditions. In these cases, knowledge or justification acquired by epistemic trust depend on social processes or social factors, because the entire belief forming process is partly social, or, more generally, non-individualist. And the properties of discrimination and defeater-sensitivity, in turn, need not depend only or even predominantly on the trusting individual. These properties can be instantiated in virtue of the social system that the subject is part of. So, an individual may be discriminate and defeater-sensitive in her trust due to social factors.

It seems apt to say that when epistemic trust is not based on evidence or reasons, it is blind. One might even suggest that genuine epistemic trust is necessarily at least partially blind. As Hardwig says: 'After all, trust, in order to be trust, must be a least partially

blind' (Hardwig 1991, 693). Epistemic trust, as I have laid out the notion, can be both discriminating and defeater-sensitive. When this is so, epistemic trust might still be blind, though fortunately not without a sense of orientation. In a recent paper Dan Sperber and others highlight the socio-epistemic importance of what they call epistemic vigilance (Sperber, Clément et al. 2010). Epistemic vigilance is the alertness to cues of lack of trustworthiness of others, or lack of plausibility of the claims they make. When meeting the conditions on discrimination and defeater-sensitivity, epistemic trust involves epistemic vigilance. As Sperber et al says: epistemic vigilance is the opposite not of trust, but of blind trust (Sperber, Clément et al. 2010, 363). What they mean by blind trust is exactly what I call indiscriminate and defeater-insensitive trust.

3.3 The problem of extended reliability

On the proposed account, epistemic trust involves a dispositional property to accept information delivered by particular sources. This dispositional property itself is a property of an individual, and not a property of the individual's social system or environment.

But have suggested two ways in which knowledge and justification acquired by epistemic trust may, and typically do, depend on factors that go beyond the individual, factors that a non-individualist in the sense of (Goldberg 2010). First, for epistemic trust to generate a known belief, one must trust a trustworthy source. In other words, there need to be a reliable process of information acquisition extending beyond the cognitive system of the trusting agent. Second, the properties of discrimination and defeater-sensitivity need not themselves be grounded in properties of the subject. Instead, they can be features of the social system that the agent is a part of.

As (Goldberg 2010) discusses, there is a particular way in which this dependence on non-individualist factors is controversial even among process reliabilists. Suppose that A trusts source I regarding the information that p (where I may be another individual, an organisation, a tradition and so on). So, A believes that p on the

ground that the source I says so, and because A trusts I on these matters. Suppose further, that A is discriminate and defeater-sensitive with regards to sources like I, and that p is indeed true. No evidence is available to A that would suggest that I is not reliable or trustworthy. However, as it happens I is not reliable, and it is mere chance and good luck that, on this occasion, A receives the information that p, rather than the information that not-p.

It seems clear that when this element of epistemic luck is brought out in the open, we don't want to say that A knows that p, although S's belief that p is true, and acquired in what to A seems to be a respectable way. But is A epistemically justified in her belief that p, and is the case consequently a Gettier Case? Or should we rather say that A is not even epistemically justified in her belief, since the relevant process by which the belief was formed is not reliable? If we chose the latter, the implication is that doxastic justification depends on social factors in a way that has not previously been acknowledged.

Goldberg argues that process reliabilists are in fact committed to accepting what he labels 'the extendedness hypothesis' (Goldberg 2010, 79ff). According to this hypothesis, the properties in virtue of which a subject's belief is epistemically justified may be non-individualist; they may involve properties of a subject's environment or the social system in which the subject is embedded because the process by which information is relayed to the subject extends spheres. If we accept the extendedness hypothesis, we should say that the subject lacks epistemic justification in the case above.

I am sympathetic to the extendedness hypothesis, though a full discussion of the issue is beyond the scope of this paper. Here I just want to note that whether one accepts the extendedness hypothesis or not, as a process reliabilist one should agree that true beliefs acquired by epistemic trust in trustworthy sources may indeed qualify as known and epistemically justified. Epistemic trust should therefore be welcomed in the process reliabilist's inventory of respectable belief forming practices. The above dispute, interesting as it is, concerns the more specific question of how to categorize

cases like the above, not whether epistemic trust is an acceptable mode of belief formation.⁹

4. Trust, Evidence and Reasons

I now want to relate the proposed view on epistemic trust to certain other views about trust and the epistemic role of trust.¹⁰ What I have proposed is, in a sense, a non-evidentialist conception of epistemic trust. Having epistemic trust in a source of information does not amount to having evidence that the trusted source is trustworthy, and neither does trusting require having such evidence.

Clearly, there are other uses of 'trust'. One can trust the truth of propositions, say when trusting that the bridge will bear when crossing. I can trust Adam to do something in part because I believe that Adam knows that I depend on him doing it. One can trust very basic propositions in the absence of evidence for them, say the proposition that there is an external world. One can trust the reliability of very basic ways of forming belief or the truth or validity of basic epistemic principles. These forms of trust take different objects (propositions, individuals, methods, principles). I would be inclined to think that they involve a variety of different psychological

⁹ In a comment to Hardwig's papers, Goldman remarks that Hardwig's notion of blind trust would seem to have skeptical implications, since trusting blindly is incompatible with rational justification (Goldman 2001, 86). However, this remark did not seem to concern the sort of process reliabilist take on epistemic trust proposed here. Of course, Goldman might object to Goldberg's claim that process reliabilists are committed to accepting the extendedness hypothesis. But even so, he should welcome epistemic trust as laid out here, even with the sort of blindness and social dependency that pertains to it. This view as such does not commit to the extendedness hypothesis.

¹⁰ One should bear in mind that we should probably not view these theories as competing theories concerning the same subject matter, but rather as attempts to conceptualise slightly different phenomena

states and relations. It seems reasonable to think that these uses of 'trust' bears some family resemblance, but I am skeptical about the idea of there being one unified phenomenon of trust which is instantiated in these different cases. But not much depends on this. I have tried to detail a specific notion of trust by specifying a functional role. Whether the form of trust thus identified bears a relation of family resemblance to what is expressed by other uses of 'trust', or whether there is some more unified account to be given, is a separate question that we need not address for the purpose of the current paper.

Some uses of 'trust' relates directly or indirectly to evidence. Elisabeth Fricker, for example, relates trust to rational deference and evidence. On her view it is rational to accept some belief on the 'basis of trust in another's testimony' only if one has, roughly, evidence for the competence and sincerity of the testifier, and 'one is not aware of significant contrary testimony regarding P' (Fricker 2006, 232).

Russell Hardin speaks about trusting others as equivalent to rationally expecting that they will honour their commitments. This just means that I can trust someone in the sense that I have good reasons for believing that the incentive structure affecting this individual is such that he will know or believe that it is in his rational self-interest to collaborate (Hardin 1996). So for Hardin, trust is very much like a doxastic attitude, the content of which we can have evidence for, and presumably trust is appropriate just when we have this evidence.

Somewhat in contrast to this, Paul Faulkner states that in 'trusting a speaker we adopt a credulous attitude. This attitude is basic and cannot be reduced to the belief that a speaker is trustworthy or reliable' (Faulkner 2007, 305). Faulkner's theory of the epistemic role of trust is elaborate and deserves a more detailed discussion. Here I just want to note two things. First, like many others, Faulkner thinks of trust as a prediction related to very particular types of social situations essentially involving my reliance on others. My trust in you amounts to my prediction that you will act in a certain way in part because I know or believe that you have

realised that I depend on your acting in just that way, and I know or believe that this will motivate you. It is not obvious how this extends to trust in say institutions or traditions, and yet there is a sense in which we can be said to trust such items. Second, on Faulkner's account, trust concerns the sincerity of an agent, but not this agent's epistemic competences. Trusting that B is sincere involves expecting that, given the appropriate occasion, B will report to me anything that she regards as true and relevant for me, and nothing she regards as false or irrelevant, within a particular domain. Clearly, this need not involve anything about B's competencies in the domain. By contrast, the property of trustworthiness, as I defined it, involves both sincerity and competence.

In an influential paper on trust, Richard Holton asserts that 'sometimes we can trust a friend to speak knowledgeably and sincerely, without believing that they will. As a result of this we will believe what they say.' (Holton 1994, 75). So, on this view, trusting is something we may resort to when we we don't believe, and don't have sufficient evidence for believing, the proposition that someone is trustworthy. So, trust is not a doxastic attitude, and does not require evidence. Holton writes: 'Trust, I will suggest, is a distinctive kind of attitude involving a distinctive state of mind.' (Holton 1994, 63), and he has interesting things to say about what kind of non-doxastic attitude trust is. By contrast, I suggest a notion of epistemic trust whereby trust involves a disposition to accept information, and this need not imply the existence of any particular type of psychological state. Trust in my sense may typically involve and causally depend emotional states like feelings of confidence, and so on, but need not. And epistemic trust in my sense does not consist of a psychological state. However, as mentioned earlier, there may be no deep disagreement here. Different views on trust might not advance competing theories of the same notion of trust, but simply specify different notions.

Holton also asserts that one can to some extent decide to trust (Holton 1994), and he offers some examples in support of this. Here might be an interesting conflict with a role of trust in epistemology.

Suppose I don't initially trust your views about some topic, which happens to be extremely important for me. Unfortunately, I am not in a position to get this information from anywhere else, so I cannot acquire the knowledge or doxastic justification I need. Can I solve this problem by simply deciding to trust you? Assuming that you are in fact trustworthy, but that I am not inclined to accept what you say, say in part because I have no evidence pertaining to your sincerity and competence. Can I nonetheless acquire knowledge or doxastic justification from by simply deciding to trust you? Intuitively, this does not seem right, so any account of trust suggesting that one might acquire knowledge or doxastic justification as a result of volitional trust would need to explain how this is possible.

My proposed account of epistemic trust does not support the idea that we can decide to trust. I cannot decide to have a disposition to believe what someone says, just as I cannot decide to believe a particular proposition. At most I can decide to *accept* what you say. I can, of course, decide to *cultivate* epistemic trust in you, or at least I could decide to *try* to cultivate a certain pattern of epistemic trust. Indeed, one might argue that an important part of our intellectual upbringing consists precisely in inculcating certain patterns of epistemic trust; through our education and training we acquire dispositions to trust certain sources of information, and not others. Similarly, by policing and design, we might try to create and maintain an environment such that the sources of information that we in fact tend to trust are just the trustworthy ones.

4.1 *Trust and epistemic reasons*

On the view I propose, my epistemic trust in a particular source of evidence is a disposition to accept as true what this source says. If we think of *epistemic reasons* as known or justifiably believed propositions in the light of which it is epistemically rational to accept certain other propositions, then epistemic trust is not an epistemic reason. Similarly, if we think of prototypes of *evidence* as perceptual states or perceived facts, then again epistemic trust is not evidence. So, while it might be perfectly legitimate to say that I believe that *p* in

part because I trust a given source, this should not be taken to imply that my epistemic trust in this source is an epistemic reason or evidence for my belief that p.

Just as epistemic trust does not provide epistemic reasons or evidence, epistemic trust does not need epistemic reasons or evidence. If one has epistemic reasons or evidence for the trustworthiness of a particular source of information, then there is a sense in which one ought to believe accordingly, and in consequence one ought to align one's confidence in this source. So, clearly, one can trust a source for reasons, and a form of trust might be a rationally required response to such reasons.

On the view I propose, epistemic trust is not evidence or epistemic reasons, and epistemic trust need not be sustained by evidence or reasons. This might invite a specific objection that I now turn to:

Suppose that Adam tells me that p. As it happens I have some evidence E against p, though not terribly strong evidence. But since I trust Adam on this matter, I assume that the evidence against p must be misleading, and I end up believing that p with a high credence. So, my epistemic trust in Adam regarding p in a sense defeats or outweighs the evidence that I have against p, or at least it makes me discount that evidence.

How, one might ask, can a purely non-evidential conception of epistemic trust account for the fact (or the putative fact) that epistemic trust interacts with epistemic reasons or evidence in these ways? How can it be epistemically rational to discount my evidence against some proposition on the ground of epistemic trust, when epistemic trust is neither evidence, nor based on evidence (and is not even a doxastic state)?¹¹

¹¹ An objection to certain views on testimony exploiting this was set forth by Jennifer Lackey, and is discussed in a draft by Paul Faulkner. Faulkner provides an interesting attempt to show 'how the

I propose a simple explanation: it is not epistemic trust as such that interacts with evidence, but the resulting beliefs acquired by trust and held with a certain level of rational credence. Suppose that Adam is in fact trustworthy on the relevant matter, and that my trust is discriminate and sensitive. I then have a high rational credence in my belief that *p*. Given this, it may be rational for me to infer that *E* is misleading. So, my epistemic trust in Adam is part of what makes me rationally discount the misleading evidence *E*. But it is not that epistemic trust itself is a separate bit of evidence that serves as an undermining or overriding defeater of *E*.

5. Concluding remarks

I have argued is that if we accept process reliabilism, there is a respectable notion of epistemic trust that may sustain knowledge and justified belief. How much this matters in practice is an open question about which I have not said much. Still, this highlights a way in which some of our knowledge and justification may be socially dependent. Sometimes our knowledge and doxastic justification depends on factors that extend beyond our own cognitive system and into the wider socio-epistemic environment of which we are a part. This holds both for the trustworthiness of sources of information, but also for the properties of defeater-sensitivity and discrimination, two properties that may in part reside in features of our wider socio-epistemic environment. In turn, this suggests an important meliorative aspect of social epistemology. We can improve the spread of knowledge and justification by designing and maintaining social systems that improve trustworthiness of the array of information-relaying sources, social systems that improve discriminative abilities (whether based on social or individual properties), and by improving sensitivity to defeaters (whether based on social or individual properties).

decision to trust can make good epistemic sense even if it is not based on the evidence' (Faulkner 2010, 5). Interesting as it is, I will not discuss the details of this here. My conception of epistemic trust is different from Faulkner's on a number of points.

References

- Adler, J. (1994). "Testimony, Trust, Knowing." Journal of Philosophy 91(5): 264-75.
- Burge, T. (1993). "Content Preservation." Philosophical Review 102(4): 457-88.
- Faulkner, P. (2007). "A Genealogy of Trust." Episteme: A Journal of Social Epistemology 4(3): 305-21.
- Faulkner, P. (2010). "The epistemic rationality of Trust." In preparation.
- Foley, R. (2001). Intellectual trust in oneself and others. Cambridge, Cambridge University Press.
- Fricker, E. (2006). Testimony and Epistemic Autonomy. The Epistemology of Testimony. J. Lackey and E. Sosa, Oxford University Press: 225-50.
- Goldberg, S. (2010). Relying on others : an essay in epistemology. Oxford, Oxford University Press.
- Goldman, A. (2001). "Experts: Which Ones Should You trust?" Philosophy and Phenomenological Research 63.
- Goldman, A. I. (1979). What Is Justified Belief. Justification and Knowledge. G. Pappas. Dordrecht, D. Reidel.
- Goldman, A. I. (1986). Epistemology and Cognition, Cambridge: Harvard Univ Pr.
- Hardin, R. (1996). "Trustworthiness." Ethics 107(1): 26-42.
- Hardwig, J. (1985). "Epistemic Dependence." Journal of Philosophy 82: 335-49.
- Hardwig, J. (1991). "The Role of Trust in Knowledge." Journal of Philosophy 88(12): 693-708.
- Holton, R. (1994). "Deciding to trust, coming to believe." Australasian Journal of Philosophy 72(1): 63-76.
- Kappel, K. (2010). On Saying that Someone Knows: Themes from Craig. Social Epistemology. D. Pritchard, A. Millar and A. Hughes, Oxford University Press.
- Sperber, D., F. Clément, et al. (2010). "Epistemic Vigilance." Mind & Language 25(4): 359-93.