

## ORIGINAL ARTICLE

# Hinge trust\*

**Annalisa Coliva** University of California, Irvine,  
California, USA**Correspondence**Annalisa Coliva, University of California,  
Irvine.Email: [a.coliva@uci.edu](mailto:a.coliva@uci.edu)

\* This paper has developed over a considerable amount of time and has been presented, in its various versions, to multiple audiences. Reasons of space prevent me from listing them all here. I wish, however, to express my gratitude to the organizers of the 2024 Rutgers Epistemology Conference for inviting me, to the official commentators for their probing questions, and to the audience for invaluable feedback. I am also grateful to people in attendance at the conference on “Trust, distrust, and forgiveness”, held at the University of California, Irvine in 2022, for very useful comments. Finally, I would like to thank Ted Mark for his generous and careful revision of the penultimate draft of this paper.

**Abstract**

Trust is central to epistemology, particularly in accounts of testimony, where it describes the relationship between a hearer and a speaker (or trustor and trustee), enabling the acquisition of information. The speaker’s trustworthiness—marked by sincerity and knowledge—is essential for testimony to transmit knowledge or justified belief. However, trust’s nature and role remain conceptually elusive, as the current debate highlights. This paper addresses the foundational question of what trust entails, rather than the conditions under which one is trustworthy. Specifically, we examine Wittgenstein’s *On Certainty* to propose a characterization of trust in its most fundamental form, termed “hinge trust.” Hinge trust is a stance preceding the ability to form justified beliefs, directed not only at people but also at perceptual faculties, objects, and the environment. It underpins our epistemic practices, particularly in acquiring epistemic hinges essential for reasoning and inquiry.

Building on this, we advocate a “trust-first” framework, analogous to the “knowledge-first” approach in epistemology. Trust is conceptualized as a primitive stance, distinct from “reliance +” a reactive attitude, or goodwill or commitment. These elements, while significant, are not constitutive of trust. Additionally, we explore

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Philosophy and Phenomenological Research* published by Wiley Periodicals LLC on behalf of Philosophy and Phenomenological Research LLC.

the interplay between trust and distrust, arguing that trust is both conceptually and axiologically prior to distrust. Finally, we address the role of trust in testimony and hinge epistemology, demonstrating its foundational significance.

## 1 | INTRODUCTION

Trust is commonly considered to enter epistemology in an account of testimony, as the relation between a hearer and a speaker, or even the trustor and the trustee, in circumstances where the former relies on the latter to acquire the information that  $p$ . In particular, the speaker needs to be trustworthy – that is, knowledgeable and sincere with respect to “ $p$ ” – in order for testimony to transmit knowledge (or justified belief).<sup>1</sup> Yet, trust remains an elusive notion, as the current debate about its nature and role shows. In fact, one might ask, when we say that someone is trustworthy – e.g., because sincere and knowledgeable with respect to “ $p$ ” – what exactly are they *worthy of*? What is that trust they are worthy of (in virtue of being sincere and knowledgeable)? The conceptual priority of trust over trustworthiness and the need to clarify what trust is are often obscured by asking when such trust is *merited* – that is, by asking under which conditions one is indeed trustworthy. Yet, in the case of testimony, that is the “easy” question, the answer to which is “When one is sincere and knowledgeable with respect to what one is saying.” Of course, in specific contexts it can be difficult to determine if one is indeed sincere and/or knowledgeable, but this is not so much a conceptual difficulty as a practical or an empirical one.<sup>2</sup> From a philosophical point of view, the conceptually difficult question is to clarify what the trust one is worthy of is, when one is trustworthy.<sup>3</sup>

By looking at Wittgenstein’s *On Certainty*, we will first offer a characterization of trust in its most basic form and will show how it enters epistemology not just when testimony is concerned, but also in the process of acquiring hinges that are essential to all our epistemic practices (§1). Indeed, by looking at the role of trust in the acquisition of hinges, we will be able to characterize it in its most basic or fundamental form. As will become clear, trust is a specific kind of stance which comes before the ability to form justified beliefs for or against empirical propositions. Such a stance may be directed not only at people but also at one’s own perceptual and cognitive faculties, objects, artifacts, and various features of one’s environment. Given the fundamental nature of such a stance and the fact that it manifests itself in the clearest form when we are considering its

<sup>1</sup> For a hinge account of testimony, see Coliva (2019).

<sup>2</sup> This difficulty arises in particular when trust in experts is concerned, especially if experts disagree over the target issue. Some theorists, moreover, have challenged the idea that being knowledgeable and sincere are necessary and/or sufficient conditions for the transmission of knowledge (or justification) via testimony. We need not dwell on these cases for here we are concerned with trust per se, rather with the conditions in which testimony transmits knowledge (or justification). There may be further conceptual quibbles over whether these are necessary and jointly sufficient conditions which need to obtain in all cases of the transmission of knowledge or justification through testimony, yet they do so in central or typical cases of testimonial transmission of knowledge and/or justification.

<sup>3</sup> The other difficult problem from an epistemological point of view is to determine the precise content and role of the assumption of someone’s trustworthiness in the structure of testimonial justification. This is the topic I took up in Coliva (2019). See also §4 of this article.

role with respect to hinges, I will call it “hinge trust.” Yet, this is just a “catch phrase.” For hinge trust – I’ll claim – is just trust and not a subspecies of it.

We will then consider the bearing of hinge trust on current debates about trust (§2) and will make a case for a “trust-first” redressing of those debates. That is, for ceasing to analyze trust as “reliance +” some other factor, such as goodwill, benevolence, commitment, etc. By analogy with what happens in the “knowledge-first” literature, we will argue in favor of taking trust to be a primitive notion, characterized as a specific kind of stance. This theory of trust can then be used to explain complex normative relations, especially between people, in which goodwill, benevolence, commitment, etc. do play a role. While these further notions can and often do combine with trust, and can justify certain reactive attitudes, neither the former nor the latter are constitutive of trust. Or so it will be claimed.

After having argued in favor of a trust-first approach, we will consider the relation between trust and distrust (§3). Doing so will offer further evidence in favor of a trust-first approach. More specifically, we will argue in favor of considering trust as prior to and axiologically superior to distrust. Once again, this can be obscured by the fact that in many real-life situations we are pre-occupied with determining the trustworthiness of other people, where the stakes are very high, such that it may be advisable to start with neither trusting nor distrusting, and to trust only once their trustworthiness has been ascertained (to a reasonable degree).

Finally, in closing (§4), we will look at the role of trust with respect to testimony and in hinge epistemology more generally.

## 2 | HINGE TRUST IN *ON CERTAINTY*

In *On Certainty*, Wittgenstein returns repeatedly to the nature and the role of trust, in the context of an epistemological investigation. There are about thirty entries in which “trust” (*Vertrauen*) and its cognates occur.<sup>4</sup> His remarks point in the direction of conceiving of trust as a basic stance of openness and reliance onto people, objects, cognitive faculties, and institutions. Trust, for him, is also the attitude we have towards “hinges” – that is, a host of propositions that have the form of empirical propositions but play a normative role in our system of judgements.<sup>5</sup> Despite their empirical form, propositions such as “There are physical objects”, or “The earth has existed for a very long time” are not up for verification; rather, they are norms of evidential significance – they need to stay put for mind-dependent evidence, such as perception, or recent findings, to accrue to a justification for ordinary empirical propositions about specific mind-independent objects, or specific historical or geological events in the distant past. Were they not taken for granted, we

<sup>4</sup> It is interesting to note that in a couple of entries where the English translation has “trust” (or its cognates), the German has “*Verlassen*,” which is sometimes translated as “reliance.” This, however, just goes to the point that in many contexts the two terms are perceived (in English and beyond) as interchangeable. Furthermore, note that in Italian while there is a specific word for “trust” and for the verb “to trust” – that is, “*fiducia*,” and “*fidarsi*” respectively – the word for “trustworthy” is “*affidabile*,” which is also how “reliable” gets expressed in Italian (only in the context of testimony, does Italian have “*attendibile*” and “*credibile*” as preferred translations of “trustworthy”). This is further testament to the fact that, from a linguistic point of view, in several modern languages it is hard to trace a sharp or relevant distinction between “trust” and “reliance” and their cognates.

<sup>5</sup> This point has been noted by several Wittgenstein scholars, including Moyal-Sharrock (2005) and Wright (2004). Pritchard (2023) is a dissenting voice but mostly due to an understanding of trust based on Wright (2004), or else deferring to moral accounts of trust (see §2), which is arguably at odds with Wittgenstein’s own position. For further discussion of this issue, see Coliva (2025).

could not consider perceptual experiences or those findings as justifications for ordinary empirical beliefs about physical objects in our surroundings or for beliefs about specific historical or geological events.

Moreover, in some cases hinge propositions function as meaning constitutive norms for Wittgenstein. For instance, if it were called into question that this, that I hold up in front of myself is a hand, then the very meaning of “hand” would no longer be clear. That is, for Wittgenstein meaning does not depend just on a uniformity in definitions but also on a uniformity in judgements – at least, in canonical or paradigmatic applications of the relevant words. Thus, in those circumstances, the judgement that this is my hand is not subject to verification and checks like a hypothesis or a genuinely empirical proposition. Rather, it functions as a standard against which my understanding of the word “hand” is measured. Given the basicness of such a stance and the fact that it is the stance we have towards hinges, it seems apposite to call it “hinge trust.”<sup>6</sup> Yet, to repeat, “hinge trust,” is neither a particular kind or (sub)species of trust, nor the kind of trust we have only towards hinges.

Hinge trust is a stance because,<sup>7</sup> as we will see, it comes before being able to have propositional attitudes towards what one trusts – at least doxastic propositional attitudes such as *j*- or *k*-apt beliefs, which require reasons in support of the propositions believed. Furthermore, it is a stance of openness in the sense that it is an *unquestioning* attitude towards what one trusts. Such an unquestioning attitude involves relying on what one trusts – which can be an object, a perceptual or cognitive faculty, a person (including oneself), a process, or an institution. As we will see, while in some cases such a stance can be initiated after having verified the trustworthiness of the trustee, from an epistemological point of view it is only by trusting first that one will be able to acquire the means to assess the trustee’s trustworthiness. Hence, trust is epistemically prior to, and independent of the determination of trustworthiness.

Indeed, Wittgenstein repeatedly makes this point in connection with hinges and insists on the fact that hinge trust – in the dual sense of basic trust and trust towards hinges – is independent from a person’s trustworthiness. For we do not trust hinges such as “There are physical objects” or “The earth has existed for a very long time” because the person who is passing them on to us is trustworthy (OC 23, 137, 671–3). Writes Wittgenstein:

Even if the most trustworthy of men assures me that he knows things are thus and so, this by itself cannot satisfy me that he does know. Only that he believes he knows. . . . The propositions, however, which Moore retails as examples of such known truths are indeed interesting. Not because anyone knows their truth, or believes he knows them, but because they all have a similar role in the system of our empirical judgments. (OC 137)

Personal trustworthiness is subject to appraisal. Such an appraisal presupposes the possession of a language and of epistemic methods apt to ascertain whether a person is trustworthy – in particular, in the epistemic case, if they are a trustworthy informant (Coliva, 2019). Hinges, however, are in turn constitutive of meaning and epistemic methods, including those that we would utilize to evaluate the epistemic (and/or deontological, and/or ethical) credentials of a person. Thus, our

<sup>6</sup> Herewith I will presuppose the account of hinges in the structure of empirical justification provided in Coliva (2015, 2022, ms).

<sup>7</sup> I am not going to provide an analysis of stances. The account of trust offered here is meant to be compatible with whatever that analysis might turn out to be – that is, in terms of dispositions, practical abilities, capacities, etc.

trust in hinges cannot depend on the fact that we have acquired them from a person we deem trustworthy. As argued in Coliva (2019), that is at most the content of an assumption, rather than of a j- or k-apt belief about the sincerity and competence of the attestor. Rather, we trust people (as well as objects, cognitive faculties, institutions, etc.) and acquire hinges from them. Yet, hinges stay put in their turn because of the peculiar role they play in inquiry. Once this machinery is in place, we can then go on and assess the trustworthiness of the attestor, if need be, or indeed the epistemic credentials of hinges.<sup>8</sup>

Moreover, for Wittgenstein, trust is manifested not only towards people, but also towards our faculties, like our senses and our memory (OC 34, 125, 133, 337). As he writes (OC 34, 337, my emphases),

If someone is taught to calculate, is he also taught that he can rely on a calculation of his teacher's? But these explanations must after all sometime come to an end. Will he also be taught that he can *trust his senses* – since he is indeed told in many cases that in such and such a special case you cannot *trust them?* – Rule and exception.

... If I do a calculation I believe, without any doubts, that the figures on the paper aren't switching of their own accord, and I also *trust my memory* the whole time, and *trust it without any reservation*. The certainty here is the same as that of my never having been on the moon.

Furthermore, according to him, we trust artifacts such as textbooks (OC 599, 600), even if we have only a rough idea of how they are produced.<sup>9</sup> Wittgenstein writes (OC 599–600, my emphases),

For example one could describe the certainty of the proposition that water boils at circa 100°C. ... The proposition is a very elementary one in our *text-books*, which *are to be trusted* in matters like this because ...

What kind of grounds have I for *trusting text-books* of experimental physics? I have no grounds for not trusting them. And I trust them. I know how such books are produced – or rather, I believe I know. I have some evidence, but it does not go very far and is of a very scattered nature. I have heard, seen and read various things.

We also trust experts, without having the means to evaluate their expertise (OC 604). Indeed, the whole point of consulting experts is to have them supply information, explanations and understanding of the target issues that we ourselves would not be capable of acquiring (for either pragmatic reasons or epistemic ones or both).

Once again, we do not trust our senses, memory, textbooks, or epistemic authorities, such as teachers and textbooks, presumably written by experts in their fields, because they have proved

<sup>8</sup> Or at least of those that in Coliva (2023, and ms.) I call “de facto” hinges. That is, those hinges, like “I have hands”, “My name is AC”, etc., which are not presupposed by all empirical inquiries but only some, or in context, and that have been “hammered” into one by continuous experience, memory, etc. For a precise characterization of this class of hinges as opposed to those that I dub “de jure” – which are presupposed by all empirical inquiries and that I consider constitutive of epistemic rationality – see Coliva (ms.).

<sup>9</sup> Compare trusting Google and Google maps in our everyday lives with generally a very rough idea of how they work. I consider trust in AI in Coliva (ms.).

trustworthy or because we deem them so. Rather, we first trust them – i.e., we unquestioningly rely on them – and that allows us to acquire epistemic methods with their characteristic hinges. In many cases, we won't have or ever develop the competence necessary to assess their trustworthiness. That's why trust in them is basic, at least at first and on many further occasions. Our attitude of trust towards these sources carries over to the hinges we acquire from them. Only once in possession of such hinges, can we then develop the competence necessary to assess the trustworthiness of these sources and, from time to time, the good standing of (at least de facto) hinges as well.

## 2.1 | The transcendental and psychological priority of hinge trust

According to Wittgenstein, hinge trust is basic, both in a transcendental (OC 150–151, 301–308, 509) and in a psychological sense (OC 159–162. Cf. Stern, 2017 and Moyal-Sharrock, 2005, chapter 9). For him, trust is basic in a transcendental sense because, as anticipated, it is what allows us to learn a language and acquire our methods of inquiry, with their own respective hinges, which then allow us to raise doubts and make further inquiries. Without this basic form of trust – hinge trust, that is – we could not acquire either a language or our methods of inquiry,<sup>10</sup> and we could not conduct inquiries or raise meaningful doubts, not even, eventually, about those hinges or sources we started off as trusting. Hinge trust, therefore, is a condition of possibility – and not just an enabling condition – for having a language and epistemic methods. As he writes (OC 150–151, 509, my emphases),

... How do I know that this colour is blue? If I don't trust myself here, why should I trust anyone else's judgment? Is there a why? Must I not begin to trust somewhere? That is to say: somewhere I must begin with not-doubting; and that is not, so to speak, hasty but excusable: *it is part of judging*.

I should like to say: Moore does not know what he asserts he knows, but it stands fast for him, as also for me; regarding it as absolutely solid is part of our *method* of doubt and enquiry.

I really want to say that a language-game is only *possible* if one trusts something (I did not say "can trust something").

And famously, in the very passages where the metaphor of hinges is introduced, he writes (OC 341–343):

That is to say, the questions that we raise and our doubts depend on the fact that some propositions are exempt from doubt, are as it were like hinges on which those turn.

That is to say, it belongs to the logic of our scientific investigations that certain things are in deed not doubted [i.e., they are trusted].

<sup>10</sup> At least, we couldn't acquire our specific natural language, with its characteristic semantics (as well as a specific phonetics, and pragmatics). For, recall, certain paradigmatic judgements are constitutive of meaning in a Wittgensteinian perspective. Note that this would be the case even if one were to buy into Chomskian linguistics and maintain that the structures of syntax and perhaps some basic semantic categories are innate.

But it isn't that the situation is like this: We just can't investigate everything, and for that reason we are forced to rest content with assumption. If I want the door to turn, the hinges must stay put.

Thus, it is part of the logic – that is the norms – of scientific investigation (and indeed of any other kind of empirical investigation) that certain propositions are *in deed* not doubted – that is, they are trusted.

As a stance, hinge trust is also part of our psychologically inbuilt way of approaching reality. That is, as children we tend to trust adults and various authorities, as well as certain worldly regularities. Our doing so – with no grounds or reasons in their favor, at least at that point – is crucial to the acquisition of language and methods of inquiry, with their own respective hinges, which then make it possible to go on investigating the epistemic credentials of our initial trust, if need be, both with respect to the source of the information and about the information itself. Here is Wittgenstein (OC 159–162, my emphases):

As children we learn facts; e.g., that every human being has a brain, and *we take them on trust*. ... I believe that I had great-grandparents, that the people who gave themselves out as my parents really were my parents, etc.

The child learns by believing the adult. *Doubt comes after belief*.<sup>11</sup>

I learned an enormous amount and *accepted it on human authority*, and then I found some things confirmed or disconfirmed by my own experience.

In general I take as true what is found in text-books, of geography for example. Why? I say: All these facts have been confirmed a hundred times over. But how do I know that? What is my evidence for it? I have a *world-picture*. Is it true or false? Above all it is the *substratum of all my enquiring and asserting*. The propositions describing it are not all equally subject to testing.

In addition, for Wittgenstein, evidence can be the cause of and/or can corroborate our trust in hinges, but it cannot epistemically ground it, for all our evidence depends on taking hinges for granted (OC 275–280, 429), either contextually, or globally.

Furthermore, evidence putatively in favor of hinges either presupposes them (in the case of *de jure* hinges) or is no more secure than the very hinges it should epistemically support (in the case of *de facto* ones). Conversely, the fact that evidence cannot speak against hinges should not be taken to epistemically support them either. For it is in the very nature of hinges that evidence putatively against them is either discounted or explained away. Just like we wouldn't revise “2+3 = 5” if, after buying two apples and three pears, we realize that we have only four pieces of fruit, similarly, we can hold on to a hinge when evidence seems to speak against it.

The best way of characterizing hinge trust, therefore, is as a basic stance of openness and reliance (OC 201–213, 508–509, 514–515, 571) on something and/or someone. As said, it is a stance because it comes before even being able to have propositional attitudes, let alone beliefs, if beliefs are taken to be propositional attitudes of acceptance of a proposition based on reasons. It is a stance of openness because it allows us to act and take in information without questioning either

<sup>11</sup> “Belief” here is non-j/k-appt belief. Thus, the term could be glossed as “trust”.

its source or its content. It is a stance of reliance on objects, people, our cognitive faculties, artifacts, and/or institutional practices to provide us with language, methods of inquiry, and other means we need to form judgements and beliefs. As we saw, moreover, it is a basic psychological stance we have as part of our psychological make-up, which serves us well.<sup>12</sup> For, to repeat, it is needed to acquire anything relevant to the entertainment of propositional contents and to their epistemic assessment.

Notice that even if hinge trust can be characterized as a form of reliance, this does not mean endorsing a reliabilist account of it. For, according to Wittgenstein, it is not because certain sources of information and methods are conducive to the formation of true beliefs that we trust them (or that our trust in them is justified, or otherwise epistemically in good standing). Rather, we act in a certain way – that is, we do trust/rely on our senses, memory, textbooks, experts, etc. That gives us a certain picture of the world – that is, a set of epistemic methods, with their characteristic hinges (OC 93–97, 162, 167, 262). Based on that, we then distinguish between what is true/false, justified/unjustified, known/unknown (OC 94). Thus, it is only by trusting *first* that we can then acquire the means to form and evaluate beliefs and their sources as reliable or unreliable. As he writes (OC 508–509; 514–515, 94, 205, my emphasis),

What can I rely on? I really want to say that a language-game is only possible if one trusts something (I did not say “can trust something”).

This statement appeared to me fundamental; if it is false, what are ‘true’ and ‘false’ any more?!

If my name is not L.W., how can I rely on what is meant by “true” and “false”?

But I did not get my picture of the world by satisfying myself of its correctness; nor do I have it because I am satisfied of its correctness. No: *it is the inherited background against which I distinguish between true and false.*

If the true is what is grounded, then the ground is not true, not yet false.

In the familiar terms of the Euthyphro contrast (adapted to our case), where “•→” means “because”, we can thus visualize the order of priority as follows:

Not (Socratic side): Trusted •→ Reliable •→ Verified

But (Euthyphronic side): Reliable •→ Verified •→ Trusted

## 2.2 | The phenomenology of hinge trust

Hinge trust has a distinctive phenomenology, characterized by feeling secure and certain (OC 217–222). This is true both of the stance and of its content, when it is a hinge. Yet, it is not this feeling that makes a hinge certain and secure. What makes a hinge certain is the role it plays in

<sup>12</sup> This claim should be compatible with possibly different accounts of how this stance is realized in human psychology. For instance, it could consist in certain dispositions, or capacities or abilities.

our system of judgements: it allows us to acquire epistemic methods and to engage in epistemic practices that give us evidence and justification for, and therefore, in some cases, knowledge of, ordinary empirical propositions. Moreover, it is these hinges that provide us with the means to doubt and inquire into ordinary empirical propositions.

Furthermore, hinge trust – in the dual sense of a basic attitude of openness and reliance, as well as of an attitude we most fundamentally have towards hinges – normally goes unspoken, and for this reason it is almost “invisible” and in a sense “elusive”. Once again, this applies to both the stance and its content, when it is a hinge. That is, we take it for granted and it normally goes without saying that there are physical objects, or that the earth has existed for a very long time, that this is my hand, that people know their own names, or that we can rely on our senses and memory. Here is Wittgenstein (OC 94–95, 397, 568, 103, my emphases),

But I did not get my picture of the world by satisfying myself of its correctness; nor do I have it because I am satisfied of its correctness. No: it is the *inherited background* against which I distinguish between true and false.

The propositions describing this world-picture might be part of a kind of mythology. And their role is like that of rules of a game; and the game can be learned *purely practically, without learning any explicit rules*.

Haven't I gone wrong and isn't Moore perfectly right? Haven't I made the elementary mistake of confusing one's thoughts with one's knowledge? Of course *I do not think to myself* “*The earth already existed for some time before my birth*”, but do I know it any the less? Don't I show that I know it by always drawing its consequences?

If one of my names were used only very rarely, then it might happen that I did not know it. *It goes without saying* that I know my name, only because, like anyone else, I use it over and over again.

And now if I were to say “It is my unshakeable conviction that etc.”, this means in the present case too that *I have not consciously arrived at the conviction* by following a particular line of thought, but that it is anchored in all my questions and answers, so anchored that I cannot touch it.

Conversely, trust tends to become phenomenologically salient either when things do not work as expected – whence distrust can ensue – or when we are confronted with something unfamiliar and unexpected; so that what we normally take for granted or goes without saying can no longer be assumed. This, once again, applies to hinges too. They usually operate in the background and go unspoken unless something out of the ordinary takes place.

Interestingly, raising doubts about hinges or indeed affirming that one knows them, when that is normally taken for granted, raises the issue of the legitimacy of our trust in them to the point of possibly annihilating it, at least within the philosophy seminar. Here is Wittgenstein again (OC 423, 481, my emphases. Cf. 500–501),

Then why don't I simply say with Moore “I know that I am in England?” Saying this is meaningful in particular circumstances, which I can imagine. *But when I utter the*

*sentence outside these circumstances, as an example to show that I can know truths of this kind with certainty, then it at once strikes me as fishy. – Ought it to?*

When one hears Moore say “I know that that’s a tree”, *one suddenly understands those who think that that has by no means been settled. The matter strikes one all at once as being unclear and blurred. It is as if Moore had put it in the wrong light.*

In sum, our trust in hinges is operating in the “background” of all our thinking, judging, and inquiring. Very rarely does it come to the foreground, and when it does, as it does when we do philosophy, it suddenly appears dubious. Of course, for Wittgenstein philosophy cannot overturn our basic certainties, but, to be sure, raising the issue of the rational legitimacy of our trust in hinges as we do in the philosophy seminar does have the effect of making us uncertain about it and its content. This phenomenological point – as disturbing as is – is entirely consistent with then finding, within the philosophy seminar, reasons to resist such an outcome. Indeed, *On Certainty* can be seen as aiming at just that. That is, philosophy cannot either epistemically ground it or overturn it.

### 3 | HINGE TRUST AND CURRENT DEBATES ABOUT TRUST

How does hinge trust connect to current debates about trust? To begin with, the current debate about trust focuses on the attitude rather than the content of it. Therefore, it is only to be expected that the connection won’t involve attention to hinges *qua* what we acquire by trusting people, objects, artifacts, institutions, our cognitive faculties, etc., in the process of acquiring a language and our methods of inquiry.

One might object that hinges are so peculiar that trust with respect to them cannot be used to understand trust in general. In favor of such a view, it might be noticed that while the hypothesis that we might be wrong about hinges would disrupt our cognitive and epistemic lives, the breach of trust in more ordinary contexts wouldn’t have such disastrous consequences. In response, it should be noted that not all hinges are on par. For instance, no disruption to our cognitive and epistemic lives was brought about by abandoning the hinge that nobody had ever been on the Moon.<sup>13</sup> In fact, quite the opposite is true – i.e., it was a great accomplishment of humankind. Conversely, in some cases the breach of trust with respect to non-hinges can have extremely serious consequences for our lives. People may never recover, or may take a long time to do so, when they suffer traumatic separations from their spouses, due to the breach of trust, or when they are affected by historical or calamitous events, like the holocaust or severe earthquakes, that breach the trust they had in their fellow humans or their environment. Thus, there is no reason to think that hinge trust – that is, trust as is typically manifested regarding hinges – cannot be used to understand trust more generally.<sup>14</sup>

Separating the issue of trust from trust in hinges is also unfortunate because it risks blurring or even hiding completely the significance of trust – indeed, hinge trust – for epistemology; indeed, doing so threatens to relegate the philosophical significance of trust, at most, to the epistemology

<sup>13</sup> This is notoriously one of Wittgenstein’s (de facto) hinges in *On Certainty*, composed between 1949 and 1951, so about two decades before the Moon-landing.

<sup>14</sup> I would like to thank Richard Gaskin for raising this objection.

of testimony, or to the determination of the conditions which need to obtain, in specific contexts, for trust to be merited – that is, for trustworthiness to obtain.

One can then note that trust in ordinary cases is somewhat voluntary and discretionary, in a way in which trust with respect to hinges isn't. Here it is important to note that these are features of the *act* of trusting, not of the stance we then have towards who or what we decide to trust (or not). Furthermore, as we have repeatedly seen, trust in (de facto) hinges can be revoked. Thus, once again, there is no reason to be suspicious of an account of trust as such that proceeds via an account of trust towards hinges.<sup>15</sup>

Be that as it may, hinge trust can play an important role in current debates about trust, which mostly take place in moral psychology. For by looking at these debates, it is quite easy to get the impression of a problematic heterogeneity (Simpson, 2012). The situation does not seem too dissimilar from the post-Gettier debate about knowledge and likewise, it may be wise to redress it by favoring a “trust-first” approach, similar – if only in spirit – to the “knowledge-first” approach proposed by Williamson (2000). For theorists seem to be interested in different things and somewhat artificially distinguish trust from reliance (Baier, 1986, Goldberg, 2020) even though such a distinction is hard to come by in ordinary discourse or even practices, where “trust” and “reliance” are often interchangeable.

Furthermore, they tend to focus on personal trust, rather than more widely on trust directed at objects, artifacts, and cognitive faculties, besides people. While this can be motivated by the fact that many theorists working on trust have been interested in its moral implications, it remains true that trust is not – at least not naturally – an attitude we have only towards people. Nor is it present in its canonical form only in our interpersonal relationships. This could be obscured by the fact that, starting with the personal case, theorists have packed into the notion of trust features, such as benevolence, commitment, as well as (motivated) reactive attitudes, that make sense solely or mostly in connection with people, thus rendering the notion of trust inapplicable to non-human entities (including AI, nowadays). Yet, such an exclusionary attitude is not grounded in our linguistic and conceptual framework, as we have just seen.

Nguyen (2022) is a noticeable exception, and indeed his account of trust as an unquestioning attitude is very similar to Wittgenstein's. Rightly, in my opinion, he takes seriously the possibility of trusting objects or artifacts. For instance, it is common to hear climbers say (and mean) that they trust their ropes, or to hear people say that they trust Google, Gmaps, etc.<sup>16</sup> Nguyen, however, purports to distinguish between relying on objects and faculties, on the one hand, and trusting them, on the other. The key idea is that one may still rely on objects or faculties even if one doesn't trust them because one has no better option, in context, than relying on what one doesn't trust. Yet, from a Wittgensteinian perspective, the real contrast would be between trusting or relying on objects or faculties, on the one hand, and simply making use of them because one has no other (or no better) option, on the other. For instance, we trust or rely on our car to start when we switch on the engine. In such a case, we are not merely taking a stab at starting it or betting that it will start (even though we may bet on its functioning well if we trust or rely on the fact that it

<sup>15</sup> We will further address this topic in §3.

<sup>16</sup> Here it may be remarked that trusting the ropes in a climber's case is a matter of having checked them and of being trained to trust them in relatively secure settings, before going on to use them in more complex and potentially dangerous ones. This is fine as far as it goes. The point I am making here is just that trust can be directed at objects, besides people. We will discuss whether trust is the default or when, for prudential (or psychological) reasons, it may not be the default in §3.

will).<sup>17</sup> However, when, in the absence of other or better options, we are forced to take our car to go somewhere, even when we know it isn't working properly, we are neither trusting nor relying on it to work properly. Rather, we are knowingly running the risk that it may not work as it is our only (or at least our best) option, given the circumstances.

To characterize personal (or affective) trust, then, theorists identify conditions which need to obtain over and above reliance. For instance, they appeal to reactive and affective attitudes, such as resentment or betrayal (Baier, 1986, Holton, 1994, Jones, 1996), as the criterion to distinguish between mere reliance on people – often called “predictive trust” – and (what by their lights is) “trust”, *simpliciter*, or “trust proper” – or, as is sometimes called, “personal” or “affective” trust. Yet, such criteria are problematic. For they seem to be neither necessary nor sufficient for trust. That is, one can trust someone to do something without resenting them, or feeling betrayed by them, for not complying with one's expectations, if the matter at hand is not particularly significant (*ditto* for other reactive or affective attitudes appealed to by theorists to distinguish between trust and reliance). Or else, one can resent them (or perhaps even feel betrayed) for not meeting one's expectations even if such expectations were not formed based on a relation of trust. For instance, after asking my daughter to boil some water prior to my arrival back home, because we will have guests for dinner and I would like to cook pasta, and while trusting her to do it (as opposed to merely predicting she will), I may find out that she didn't. It would be weird to think that I would thereby resent her, or even feel betrayed by her, for not doing what I trusted her to do, or else that I did not trust her in the first place. After all, boiling a pot of water takes about ten minutes and it doesn't make a huge difference if dinner is served at 7 p.m. or ten minutes later. Conversely, a demanding parent can resent their child for not meeting their academic expectations, or even feel betrayed by that, even if such an expectation was clearly not formed based on a relation of trust. That is, it would be extremely weird – and a sign of a deeply problematic relationship with one's kids – to say that one was trusting or relying on one's kids to be academically successful. At most, one might have had that expectation or hope.

What this teaches us in general is that the project of analyzing trust as “reliance +” where the further element of the equation would be a reactive attitude, is bankrupt. It thus seems wise to move away from such a strategy and instead to take trust to be a basic, *sui generis* stance that is characterized in terms of an unquestioning reliance; a reliance which we may have towards objects and people, in a variety of circumstances, independently of whether the breach of trust gives rise to reactive attitudes.<sup>18</sup> Of course these reactive attitudes may be present with respect to people and possibly even objects: after all, we may indeed resent objects for not complying with our expectations if we were relying on/trusting them to function well. Yet, these reactive attitudes should not be considered constitutive of trust, or even of specific species of trust, which would go beyond “hinge trust” (as we shall presently see).

Some theorists identify trust with trustworthiness (Hardin, 2002, Hieronymi, 2008, Hawley, 2019) – an identification that makes sense mostly if trust is limited to personal trust. Theorists,

<sup>17</sup> Notice, however, that I don't consider betting as invariably correlated with trusting/relying on someone/something. For one may indeed bet on an outcome even when not trusting/relying on someone/something, just for the sake of betting or for the projected gain of an unexpected win. Nor is trust rational only if one has calculated the odds of a positive outcome. Indeed, such calculations are possible only by possessing the relevant epistemic methods, with their attendant hinges, which we can acquire only by trusting first. Also, it remains that accounts of predictive trust are silent about trust itself. At most, they merely impose conditions on when trust – whatever it might turn out to be – might be rational.

<sup>18</sup> This is one further reason not to follow Nguyen in characterizing the difference between reliance and trust in objects as dependent on the kind of reactive attitude their malfunctioning elicits from us.

moreover, often impose demanding conditions on trust proper, which tend to precisify the notion in a way that is open to many counterexamples. That is, situations in which we would naturally speak of trust, but where the precisification of trust under consideration makes talk of trust illegitimate.<sup>19</sup> Alternatively, as mentioned, they take themselves to be illuminating “trust proper”, or the only philosophically rich notion of trust worth investigating. Yet, this seemingly less radical move is not without problems. For, if there is both “trust proper” – call it TRUST (with capital letters) – and something which is trust but not quite TRUST – call it “trust -” (to be read: “trust minus”) –, then either these are two different concepts, or they are different species of one and the same concept. If they are two different concepts, then calling both “trust” is just a recipe for confusion, for they would no more have something in common than a bank and a bank of a river. If, in contrast, they are both different species of the same concept as a genus – that is, TRUST (in small caps to indicate a concept rather than a word) –, then the question remains as to how we should characterize the concept which is their common genus. If it is just reliance, then this goes against the claim that reliance isn’t sufficient for trust.

The notion of hinge trust as a stance of openness and reliance on something/someone doesn’t call for a problematic distinction between reliance and trust; or between predictive trust and personal/affective trust, where the former is just reliance and the latter is trust proper (*contra* Baier, 1986, and Hertzberg, 1988). In particular, hinge trust is not a matter of making predictions about the behavior of people, objects, faculties, artifacts, etc., but of unquestioningly relying on them.

The notion of hinge trust as a stance of openness and reliance on something/someone (else) makes it possible to then add more detailed conditions, which may jointly constitute thicker normative notions. For instance, I may be trusting you to  $\varphi$ , because you gave me your word, or you promised me that you would  $\varphi$ ; or I may trust you to fulfill your commitment, which may have been incurred voluntarily, or just in virtue of occupying a certain social role (Hawley, 2019); alternatively, I may trust you because of your good will towards me, or in virtue of our wedding vows, etc. These further specifications would tell us something about what being a word-keeper, a reliable person vis-à-vis one’s social or professional commitments, or being a loyal friend or a faithful spouse amounts to. For instance, it may consist, at least in part, in having the other party unquestioningly rely on your doing thus-and-so, so that they can take it for granted that you will do what you said you would do, in virtue of having given your word for it. Similarly, being a trustworthy person vis-à-vis one’s commitments will consist in having others unquestioningly rely on you for the fulfillment of those commitments (at least *ceteris paribus*), either because of your promising to do certain things or because doing them is part of the role you have (e.g., in the workplace, as a parent, etc.). Being a faithful spouse, in turn, will consist, at least in part and at first approximation, in having your partner unquestioningly rely on you with respect to the content of your wedding vows, based on the feelings that should have generated or accompanied them in the first place, and out of the commitments undertaken by making those vows.

The details don’t matter much for present purposes. These examples only tell us something about the further normative relations of which trust is a component. Yet none of the further conditions which combine with trust in these thicker normative relations are constitutive of trust *per se*. That is, all these further thicker normative relations – such as being a word-keeper, a trustworthy person with respect to one’s commitments, a loyal friend and a faithful spouse, etc. – do involve trust but are not themselves constitutive of it, nor are the further conditions – such as benevolence, commitment, relevant feelings, etc. – which combine with trust to give rise to these thicker normative relations.

<sup>19</sup> Some of the conditions that are often problematically presented as constitutive of trust will be listed shortly.

The key move is thus to take trust – characterized as hinge trust – as basic and then use it to characterize other notions in the vicinity such as being a trustworthy colleague, a loyal friend, a faithful spouse, a word-keeper, etc. just like, in the “knowledge-first” approach, knowledge is taken as basic and then used to characterize other notions such as justification, reasonableness, rational exculpation, etc.<sup>20</sup>

As remarked, hinge trust is a stance, which does not constitutively involve doxastic propositional attitudes, which, in their turn, constitutively depend on reasons, such as *j*- or *k*-apt belief. It does not constitutively involve the belief that something or someone will do what we trust them to do. This is important because it allows us to characterize infants’ “thoroughgoing dependence on their parents as a paradigm kind of trust” (Simpson, 2012: 559). Moreover, belief is rational only if backed up by reasons, but trust, at least trust towards hinges, is prior to the very possibility of offering reasons. This, however, is compatible with then (or sometimes) forming such a belief. For instance, if I trust that the floor will not disappear into the abyss, I will normally just unquestioningly rely on it and act accordingly. However, if I have the concepts necessary to consider the issue and I do consider the issue, I may form the belief that the floor will not disappear into the abyss and offer evidence in support of it. Or else, if I trust you to do something for me, I will unquestioningly rely on it. Yet, if I have the concepts necessary to consider the issue and I do consider the issue, I may form the belief that you will unquestioningly do such-and-such for me, and support this with evidence. Yet these beliefs are not themselves necessary or sufficient conditions for trust.

#### 4 | THE ONTOLOGICAL AND AXIOLOGICAL PRIORITY OF (HINGE) TRUST OVER DISTRUST

What hinge trust brings to light is the Janus-faced nature of trust. On the one hand, trust is characterized by a phenomenology of feeling secure and at ease with our human and non-human environment. On the other, it is a stance that, by involving the dependance on others and on objects and artifacts to do certain things for us, constitutively opens us up to the possibility of being let down. Far from being a problem, however, this just shows that this form of dependance is in fact a condition of possibility of our success as individuals and as a species. In this sense, hinge trust illuminates what it means to say that humans are a social species. For at the heart of our individual and collective success there is a reliance on others (as well as on several aspects of our environment) that allows us, as individuals and as a species, to acquire and transmit all necessary elements for forming beliefs and assessing them. To repeat it, we don’t trust because it has proved successful. Indeed, if that were the correct genealogy of trust, we would long have been extinct before having enough inductive evidence in favor of trusting one another. Rather, we trust and that is what enables us to be successful – it allows us to possess a language and methods of inquiry that can vastly extend our knowledge and then it allows us to pass such knowledge (including linguistic knowledge) on to others. As we saw, eventually this may enable us to assess the epistemic trustworthiness of those sources from which we acquired those methods of inquiry

<sup>20</sup> This also accounts for the idea that sometimes we trust someone *because* they have promised to do something for us, or because they have made certain vows, or because they have undertaken certain commitments. Being a trustworthy colleague, a word-keeper, or a loyal spouse, as we saw, all involve trust, even though trust itself is more fundamental than any of these further normative relations one enters in virtue of promising, undertaking commitments, marrying, etc.

with their characteristic hinges and even the epistemic credentials of hinges themselves (at least *de facto* ones), but none of that would be possible without trusting *first*.

Because such a reliance on others and on several aspects of our environment is “such a routine part of life, we very often trust without talking about it. Most of the time, it just happens” (Simpson, 2012: 560). Indeed, as Simpson remarks, “the actuality of trust may be very present, but it does not need to be talked about unless there is some problem, and so trust is invisible” (*ibid.*). As we saw in §2, it is only when trust is broken, or we are facing something unfamiliar that trust surfaces in the conversation (similarly to what happens with hinges).

Being such a basic stance, trust is the default. *Distrust*, for us, is like illness with respect to health: it is *not* our “normal” condition, and it is rational or justified only when the initial conditions for trust have been systematically or egregiously violated. Thus, hinge trust is both ontologically and axiologically prior to distrust.

Now, the quantity of literature on distrust is modest in comparison to the quantity of literature on trust. In it, however, the point is often made that trust and distrust are contraries, not contradictions. That is, they are not exhaustive, for one may not trust someone without thereby distrusting them, but they are mutually exclusive. That is, if one trusts someone, then that is incompatible with distrusting them (at least in a given area) and *vice versa*.<sup>21</sup> Given that hinge trust is ontologically and axiologically superior to distrust, I remain doubtful of those theories, like Hawley’s, that aim to shed light on trust by elucidating distrust.<sup>22</sup>

Yet, it is important to clarify when distrust is rationally legitimate. As we anticipated, there is no quick and ready way of specifying when trust is systematically or egregiously violated. As such, there is no quick and ready way for establishing when moving on to distrust is rational. As a rule of thumb, it may depend on our trust either being betrayed on multiple occasions and/or on significant issues, or indeed in circumstances where we would routinely rely on people, objects, our faculties, etc. For instance, if being faithful is at the core of one’s marital relationship, being betrayed even just once may be a good reason for distrusting one’s spouse and even for ending the marriage. Or else, if one can keep trusting after being betrayed once, it may be reasonable to move on to distrust if episodes repeat. Interestingly, from an epistemic point of view, if my sight and memory, which I usually trust, let me down in what are generally considered environmentally and subjectively normal conditions – when the lighting conditions are good, or when I am not intoxicated or affected by the posthumous effects of anesthesia – then that rationally entitles me to distrust them. Whereas, if they let me down in environmentally or subjectively abnormal conditions – when the room is too dark to see, or I am recovering from a concussion, say – then it would not be rational to move on to distrusting these faculties in general. Their occasional failure is entirely compatible with the fact that they remain trustworthy sources of information or of its retention, in the general run of cases.

Again, if in what seem to be normal conditions a person or even an expert gives me incorrect information about topics in their purview, this makes it reasonable for me to stop trusting them or even moving on to distrusting them (as sources of information at least in the area of discourse under consideration). By contrast, if the topic is not one about which they can be presumed to have knowledge or expertise, or if the conditions aren’t normal (e.g., they are intoxicated, under the effect of drugs, have just been injured, etc.), then their passing on incorrect information is not

<sup>21</sup> See Hawley (2014, 2019) and D’Cruz (2020).

<sup>22</sup> Whereas the important theme of when distrust is merited or not, especially when driven by identity prejudice against someone (see Medina 2020, Scheman 2020), is partly taken up in Coliva (*ms.*).

sufficient to make it rational for me to stop trusting them or for me to start distrusting them in general and/or with respect to their area of expertise.

To repeat, since trust is the default for us, it has “ontological priority” over distrust and it is “axiologically” superior to distrust (see also Stern, 2017). For it is the stance that allows us to operate best from an epistemic point of view (and beyond). This explains why, as a stance, trust can be maintained in the face of counterevidence, at least to a reasonable degree. That is, as Descartes famously states in the *First Meditation*, it is not justified to move on to distrust if our faculties, other people, institutions, or our environmental conditions have betrayed us once or only on occasion, or if trust has been broken on non-fundamental issues, or in abnormal conditions.

#### 4.1 | (Dis-)trusting for a reason

There may be cases, however, where while distrust cannot and should not be the default, the neutral position of neither trusting nor distrusting would be the rational one to occupy. Suppose you want to hire a babysitter for your 1-year-old child.<sup>23</sup> The stakes are high and while you have no reason to distrust this person, you may start off in a neutral position, and, by considering their credentials and references first, go on to determine whether they are trustworthy. And trust – i.e., unquestioning reliance – may ensue only once such a check has been satisfactorily completed. Notice that in such a case what is at stake is the assessment of someone’s trustworthiness such that we can then hold a stance of unquestioning reliance on them with respect to taking care of our child. In such a case trust is not the default – nor is distrust, for that matter – yet the establishment of trust is the aim of the process, which may be repeated until we do find someone that we deem worthy of trust.<sup>24</sup> Furthermore, we have already considered the fact that trust can easily be broken, and that this can have disastrous consequences when stakes are high. Now, in the case under consideration it seems *prudentially* better to prevent what could be a disastrous outcome by starting off in the neutral position and then finding one’s way forward towards trust. Yet, these are prudential considerations, not epistemic ones. Indeed, we may have no good epistemic reason to doubt the trustworthiness of the person in front of us. Yet, the stakes are high, and prudential considerations may outrun the lack of epistemic reasons for distrust or for being open-minded.

That this may happen in high-stakes scenarios is in no way an argument against the fact that, in the normal run of cases, trust is the default, not just from an epistemic point of view but also from a personal one. We normally trust that unknown people on the street won’t suddenly turn into killers; we trust a taxi driver upon our first arrival in a new country to take us where we ask them to; we trust that the new colleagues we have just hired will behave collegially.<sup>25</sup> Of course, this trust can be defeated, but – once again – the fact that sometimes it is defeated should not obscure the fact that most of the time it isn’t. Nor should it be taken to show that we in fact ought

<sup>23</sup> Many thanks to Jason D’Cruz for raising this issue in conversation. Thanks also to Joost Ziff for raising a similar worry with respect to trust in romantic relationships, where the consequences of trusting someone who isn’t deserving of it would likely be psychologically disastrous.

<sup>24</sup> The courtroom case is different. While the stakes are high, one is/must be presumed innocent until proven guilty. Indeed, one might say that because the stakes are extremely high – including limitation of personal freedom if not capital punishment –, to protect the defendant from the possibility of calamitous mistakes they are presumed innocent until proved guilty beyond any possible doubt. Thus, the default is certainly not distrust *nor* the neutral position. For the presumption of innocence is incompatible with a stance of open-mindedness about it.

<sup>25</sup> Note that it would be incorrect to claim that we trust in such cases only insofar as we have amassed enough inductive evidence. Rather, we trust as a default, and revise only when (enough) contrary evidence is obtained.

not to trust until we have reached a sufficiently high level of confidence with respect to people's behavior in those circumstances. This is neither what we do, nor what we ought to do. Indeed, if we did this, it would be a symptom of psychosis. Finally, the fact that this trust can be defeated should not be taken as an argument for doubting that when trust is the result of an inquiry into the trustworthiness of someone (or something), the checking is only possible if we possess the epistemic methods needed to conduct it as well as their attendant hinges, which we have acquired through trust and are the object of our trust.

Still, there may be other cases where it is epistemically rational not to trust, even though trusting may generally be the default. For instance, we typically unquestioningly rely on experts. When we go to see our doctor for the nagging symptoms of the flu, we take their advice on trust. Similarly, we trust our mechanic's opinion about the state of our car, and what kind of service it needs. Yet, one may be getting contrasting opinions from experts.<sup>26</sup> Experts' disagreement may thus be a good reason to occupy a neutral position, while waiting for experts (or most of them) to eventually converge onto specific guidelines or a shared opinion.<sup>27</sup>

While these are certainly legitimate cases in which it is rational not to trust, they do not show that in the general run of cases trust isn't or shouldn't be the default. Nor are they incompatible with the claim we have been making all along; namely, that while trust comes first, as it allows us to acquire the means needed to make epistemic evaluations, once we have these means we can inquire further into the trustworthiness of those who passed them onto us in the first place – that is their competence and sincerity – as well as into the good standing of the hinges we have acquired from them. After all, Galileo had inherited the *de facto* hinge that the Earth was at the center of the universe, and he did develop reasons (albeit far from conclusive ones at the time) that led him to first demote it to the role of a hypothesis and then to abandon it altogether.<sup>28</sup> Those very reasons presumably made him also question the competence of those from whom he had received that hinge. Thus, none of these cases show that trust isn't or shouldn't be the default, at least in the general run of cases.

Another worry is that if trust is the default, then one may be gullible.<sup>29</sup> Yet, for several reasons, this doesn't really follow. First, in the case of children, there is no other option for them but trust if they are to acquire the means necessary to engage in epistemic practices, including those by means of which the trustworthiness of attesters is assessed. Thus, their trust is not a sign of gullibility. For gullibility consists in (culpably) ignoring or playing down the relevance of the cues that should be taken to show that people or other aspects of one's environment aren't trustworthy. Second, to say that trust is the default does not mean that it is never rational to move on to distrust or to adopt a neutral position (for reasons which may be epistemic or prudential, as we saw). Gullibility would ensue only if, in these or similar circumstances, we went on trusting in the face of excellent reasons not to. In fact, what children are typically lacking is the ability to correctly assess whether those people or even artifacts (such as the internet nowadays) they trust are indeed trustworthy, because they are not yet sufficiently responsive to (potential) defeaters.

<sup>26</sup> We were all exposed to that recently, at the inception of the pandemic, especially with respect to the usefulness of wearing masks in public.

<sup>27</sup> Experts' disagreement, however, is not a sufficient reason for distrusting experts altogether, since in many cases it may be due to lack of sufficient data or of robust explanations of them.

<sup>28</sup> Notoriously Galileo relied on tides as a reason for thinking that the Earth was rotating around the Sun when in fact tides are not (mainly) caused by the Earth's rotation, but – simplifying a lot – by the gravitational attraction exercised by the Moon (and the Sun) over the surface of the sea.

<sup>29</sup> Thanks to Maria Baghramian for raising this objection.

Finally, given the axiological priority of trust over distrust, the breaking of trust may have disastrous consequences. For turning to distrust can deeply affect the way people go through life and their being in the world. In fact, as stances towards life – that is, as ways of being in the world –, trust and distrust may be appealed to in order to mark the difference between the world of the “happy” and the world of the “unhappy” (Wittgenstein, 1921, 6.43). Whereas the world of the happy is one of openness, confidence, and ease with respect to others, the environment and even oneself, the world of the unhappy is one of closure, of insularity and lack of reliance on others, one’s environment and possibly even one’s own faculties. Distrust is therefore deeply dehumanizing, and, conversely, trust appears as a core feature of human well-being.

## 5 | HINGE TRUST, HINGE EPISTEMOLOGY AND TESTIMONIAL HINGES

Hinge trust is a key element of hinge epistemology *tout court*. For it is the stance that characterizes at the most basic level our attitude towards hinges whether they are hinges of empirical or of social epistemic rationality. Since hinge trust, by its very nature, is not supported by reasons, nor could or need be, this also indirectly speaks in favor of those forms of hinge epistemology, like the constitutive version of hinge epistemology defended in Coliva (2015), that dispense with any epistemic support for hinges. This contrasts for instance with Wright’s (2004) version of hinge epistemology, which takes trust to be our fundamental attitude towards hinges but then appeals to entitlements – that is, non-evidential warrants – to make it rational (or at least, to reflectively redeem its rationality).<sup>30</sup> For, according to Wright, trust involves risk and warrant is needed to make taking that risk rational. The key point of hinge trust, however, is that no such rational support is necessary, and it isn’t necessary precisely because it involves no risk, but just dependence on our human and non-human environment.

Furthermore, on my account of testimony (Coliva, 2019), testimonial justification depends on the testimonial hinge that people are generally trustworthy – that is, are knowledgeable and sincere with respect to what they say.<sup>31</sup> In fact, according to that account, that hinge is a constitutive element of social epistemic rationality. As such, it figures as a groundless assumption in the abstract space of reasons,<sup>32</sup> and, when no defeaters occur, it combines with what a speaker is saying to provide a listener with a testimonial justification to believe what they are told.

This, however, is the theoretical characterization of what happens in the abstract space of reasons and of how testimonial justification is possible only if the testimonial hinge is taken for granted. Yet, what in fact happens is that by trusting speakers<sup>33</sup> – that is, by having a stance of openness and unquestioning reliance on their sincerity and competence (with respect to what they are telling us) –, absent defeaters, we acquire testimonial justification for (and sometimes even knowledge of) what we are told. The notion of hinge trust, as we have characterized it, clarifies what trusting speakers consists in. Namely, it consists in a stance of openness and unquestioning reliance on their sincerity and competence with respect to the information they are imparting

<sup>30</sup> I criticize Wright’s account of entitlements in Coliva (2015, chapters 2 and 4).

<sup>31</sup> More precisely, the testimonial hinge concerns the trustworthiness of the specific attestor – that is, their sincerity and their being knowledgeable or at least justified – with respect to the specific proposition they are asserting, as claimed in Coliva (2019).

<sup>32</sup> That is, it is part of the structure of propositional justifications.

<sup>33</sup> This is typically classified as a form of objectual trust.

to us. It is such a stance, however unsupported by evidential or non-evidential reasons it may be, that, absent defeaters, is constitutive of social epistemic rationality, for it makes it possible to acquire testimonial justifications in the first place. Of course, sometimes people lie, or assert what they aren't knowledgeable about (or have no justification for). Thus, testimonial knowledge and/or justification are easily defeated. Yet, once again, this should not be taken to show that they can only be obtained after verifying the good standing of that assumption.

Thus, paraphrasing Wittgenstein (OC 253), we may say that at the foundation of all well-founded beliefs acquired through testimony lies unfounded trust – that is, a non-epistemic and even non-propositional stance of openness and reliance towards others. If, as Moyal-Sharrock (2005) argues, Wittgenstein's *On Certainty* is responsible for “bringing the animal back into epistemology”, it is with the basic stance of openness and reliance that characterizes hinge trust that the animal is brought back into the epistemology of testimony. That is, it is only by occupying such an unreflective stance that – at least at the most basic level – we can then acquire reasons of a testimonial nature for at least some of our beliefs and therefore deploy social epistemic rationality.

Finally, since all hinges – constitutive of epistemic rationality, social, or general – are acquired by trusting others and the world, the innovative aspect of hinge epistemology – social and general – is that it is not ratiocination that makes it possible for us to participate in epistemic practices wherein epistemic rationality unfolds but rather occupying such an unreflective stance. Thus, Wittgenstein's revolutionary message is not so much to have focused on the importance and primacy of action – Hume had already noticed that, outside the philosophy seminar, we act with a certainty that knows no doubt even if it is not and cannot be supported by reasons. Rather, it is to have brought to light the fact that our animal, unreflective stance of openness and reliance upon others, on our cognitive faculties, and the world is what allows us to acquire the hinges that are constitutive of epistemic rationality – empirical and social. That is, rationality and reasons would not – logically, and not just causally and genealogically – be possible without the animal; at least for finite and social animals like us.

## ORCID

Annalisa Coliva  <https://orcid.org/0000-0001-6258-5276>

## REFERENCES

- Baier, A. (1986). Trust and antitrust. *Ethics*, 96, 231–260. <https://doi.org/10.1086/292745>
- Coliva, A. (2015). *Extended rationality. A hinge epistemology*. Palgrave.
- Coliva, A. (2019). Testimonial hinges. *Philosophical Issues*, 29, 53–68. <https://doi.org/10.1111/phils.12140>
- Coliva, A. (2022). Wittgenstein rehinged. The significance of *on certainty* for contemporary epistemology. Anthem.
- Coliva, A. (2023). Hinges in the knowledge economy. On Greco's common and procedural knowledge. *Synthese*, 201, 1–18. <https://doi.org/10.1007/s11229-023-04130-5>
- Coliva, A. (2025). *Wittgenstein and social epistemology*. Cambridge University Press.
- Coliva, A. ms. *Social and applied hinge epistemology*.
- D'Cruz, J. (2020). Trust and distrust. In J. Simon (Ed.), *The Routledge Handbook of Trust and Philosophy* (pp. 41–51). Routledge.
- Goldberg, S. (2020). Trust and reliance. In J. Simon (Ed.), *The Routledge Handbook of Trust and Philosophy* (pp. 97–108). Routledge.
- Hardin, R. (2002). *Trust and trustworthiness*. Russell Sage Foundation.
- Hawley, K. (2014). Trust, distrust and commitment. *Noûs*, 48, 1–20. <https://doi.org/10.1111/nous.12000>
- Hawley, K. (2019). *Trust and distrust*. Oxford University Press.
- Hertzberg, L. (1988). On the attitude of trust. *Inquiry*, 31, 307–322. <https://doi.org/10.1080/00201748808602157>

- Hieronymi, P. (2008). The reasons of trust. *Australasian Journal of Philosophy*, 86, 213–236. <https://doi.org/10.1080/00048400801886496>
- Holton, R. (1994). Deciding to trust, coming to believe. *Australasian Journal of Philosophy*, 72, 63–76. <https://doi.org/10.1080/00048409412345881>
- Jones, K. (1996). Trust as an affective attitude. *Ethics*, 107, 4–25. <https://doi.org/10.1086/233694>
- Medina, J. (2020). Trust and epistemic injustice. In J. Simon (Ed.), *The Routledge Handbook of Trust and Philosophy* (pp. 52–63). Routledge.
- Moyal-Sharrock, D. (2005). *Understanding wittgenstein's on certainty*. Palgrave Macmillan.
- Nguyen, C. Thi. (2022) Trust as an unquestioning attitude. *Oxford Studies in Epistemology*, 7, 214–244. <https://doi.org/10.1093/oso/9780192868978.003.0007>
- Pritchard, D. (2023). Hinge commitments and trust. *Synthese*, 202, 1–20. <https://doi.org/10.1007/s11229-023-04378-x>
- Scheman, N. (2020). Trust and trustworthiness. In J. Simon (Ed.), *The Routledge Handbook of Trust and Philosophy* (pp. 28–40). Routledge.
- Simpson, T. (2012). What is trust?. *Pacific Philosophical Quarterly*, 93, 550–569. <https://doi.org/10.1111/j.1468-0114.2012.01438.x>
- Stern, R. (2017). ‘Trust is basic’. Løgstrup on the priority of trust. In P. Faulkner & T. Simpson (Eds.), *The Philosophy of Trust* (pp. 272–293). Oxford University Press.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press.
- Wittgenstein, L. (1921). *Tractatus logico-philosophicus*. Routledge and Kegan Paul.
- Wittgenstein, L. (1969). *On certainty*. Blackwell.
- Wright, C. (2004). Warrant for nothing (and foundations for free)? *The Supplementary Volume of the Aristotelian Society*, 78, 167–212. <https://doi.org/10.1111/j.0309-7013.2004.00121.x>

**How to cite this article:** Coliva, A. (2025). Hinge trust\*. *Philosophy and Phenomenological Research*, 110, 939–958. <https://doi.org/10.1111/phpr.13142>